Merlin: a decision support system for outdoor leisure planning : development and test of a rule-based microsimulation model for the evaluation of alternative scenarios and planning options

MERLIN:
A decision support system for outdoor leisure planning

Development and test of a rule-based microsimulation model for the evaluation of alternative scenarios and planning options

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de Rector Magnificus, prof.dr. R.A. van Santen, voor een commissie aangewezen door het College voor Promoties in het openbaar te verdedigen op maandag 10 december 2001 om 16.00 uur

door

Manon van Middelkoop

geboren te Voorburg

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr. H.J.P. Timmermans
en
prof.dr. J.A. Mazanec

Technische Universiteit Eindhoven
Faculteit Bouwkunde, Capaciteitsgroep Stedebouw

This research project was funded by the Netherlands Organisation for Scientific Research (NWO) as part of the FUTRO-field that was set up in 1992 to answer the need for fundamental research on time-space dynamics in recreation and tourism within the European context

Printed by the Eindhoven University Press Facilities

Cover design by Ton van Gennip, Tekenstudio Faculteit Bouwkunde

Cover photograph: Merlin's cave and the ruins of a 12th century castle generally known as "King Arthur's Castle", Tintagel, Cornwall (UK)

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG
ISBN 90-6814-560-6

Preface

According to an esteemed fellow member of the Urban Planning Group, writing the preface of one's thesis is the hardest part. Theo, I disagree with you about this "*stelling*", but thank you for providing me with the first sentence!

More seriously, writing this preface does mean that a rather memorable chapter of my life will soon come to a conclusion. This is the time to put those who contributed to this thesis in the limelight. In 1996, Harry Timmermans, "welcomed me aboard" the Urban Planning Group because during the interview we had discussed my family's interest in aviation and his numerous trips to Canada and the USA. This is a typical example of his readiness and ability to develop a language that links up with his students. Many examples have followed since, and I sincerely thank him for being my academic "captain" over the past five years.

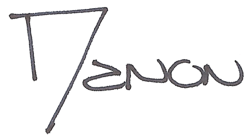
Those who wish to thank Harry cannot get round the women that support him. Ria, his wife, is the best First Lady any member of the Urban Planning Group could wish for. Mandy van de Sande-van Kasteren is the Group's secretary and the pivot around which everything revolves. I have really enjoyed her enthusiasm, her assistance and her sportsmanship! Astrid Kemperman, finally, who started off as a Ph.D.-student but who now is an established member of the Urban Planning Group, has always been my "partner in recreation and tourism" within the Group.

In addition, Harry has surrounded himself with several remarkable men. In those rare cases that Harry's ability to speak my language would fail him, Aloys Borgers was always there to take over. Besides simply being a very nice person, Aloys is especially valued for meticulously going through each paper we have produced together. Theo Arentze, Peter van der Waerden, Han Lörzing and Leo van Veghel complement Harry's crew each in their own special way.

Outside Eindhoven I am especially indebted to my second advisor, Professor Josef Mazanec at the Viennese Institute for Tourism & Leisure Studies. He has contributed to this thesis both directly and indirectly. Directly, his detailed evaluation of the final manuscript is much appreciated. More indirectly, he was closely associated with organising two international meetings that have influenced my work and thinking tremendously.

Back in Eindhoven, the social structure of the Urban Planning Group rests on four pillars. Without this structure my work would have been much less agreeable and I shall therefore share the related house rules with you. Don't ever try to call the Group on Thursdays between 10 and 11 AM, because during this coffee break the "*vlagenda*" rules and even the otherwise very complaisant Mandy will not see you. Second, on Mondays, Wednesdays and Fridays, you run the risk of being persuaded into joining the activities of the "sports group" at lunchtime. Third, newly arrived PhD-students cannot refuse to attend the notorious "*AIO-borrels*" (former students usually cannot resist). And finally, there are always one or more roommates who care for another cup of tea and who are there when you need to let off steam. During the first three years on the 11th floor, physically and psychologically separated from the rest of the Group on the 5th floor, the company of roommates was even more important, and I thank Martijn Klabbers for the many cups of tears and tea and laughter we have shared there.

Unfortunately, the path leading towards this thesis has not been strewn with roses all the way. In October 1997, the first signs of repetitive strain injury surfaced and life has not been the same since. I owe this thesis to those who pulled me through this ordeal. In particular to Perry Broers, the physiotherapist who helped me back on my feet again, physically as well as mentally. Also to Dave Janssen, who, as part of his period of practical training, put life into MERCIN by programming the system in a very admirable way. To my family and friends and colleagues who would always lend a helping hand or ear. And most importantly, to Ger Kwakkel, my friend and counterbalance for 10 years now, who took care of me whenever I was incapable of doing so myself. Thanks!



Eindhoven, October 2001

Table of Contents

List of Tables

List of Figures

1 Introduction

This introductory chapter contours progress in outdoor leisure planning in the Netherlands over the last four decades and identifies the need for a new approach to modelling outdoor leisure behaviour. Based on this review the aims and objectives and the research questions underlying this study are explicated, and the organisation of this dissertation is outlined.

1.1 Outdoor leisure planning in the Netherlands

People have been always involved in leisure and travel activities. It was not until after the Second World War, however, that, propelled by increased leisure time and economic wealth, mass outdoor recreation and tourism started to grow steadily. In response to the increasing consumption of leisure goods and services, the first government policies and planning tools were developed to facilitate the growth of recreation and tourism in the 1950s and 1960s. At the same time, scientists and professionals started to systematically analyse the phenomenon. Many theories and models of preference and choice have been advanced in order to forecast the potential impacts of policy and planning decisions. Typically, shifts in the focus of recreation and tourism planning concepts and policies were accompanied by new developments in behavioural theories and models.

During the 19th century, Europe transformed from an economy based on agriculture to one based on industry and commerce. A process of massive urbanisation accompanied the industrialisation as agricultural practice forced people to leave rural areas, and newly established factories drew them to the rapidly expanding cities. The living conditions in these cities were poor, but demand for recreation and tourism was low because the general public could not afford it.

By the 1930s, however, rising income, increased leisure time, paid holidays and increased mobility had instigated a substantial call for public open space in and near cities[1]. A new cultural philosophy emerged, emphasising the compensating function of leisure and advocating the idea that leisure, outdoor recreation in particular, provided a necessary escape away from the city. These changing attitudes paved the way for the establishment of the first outdoor leisure facilities and areas that reflected a great appreciation of nature, sobriety, simplicity and independence. Examples of these first planning concepts for recreation include the “Amsterdamse Bos” Park in the Netherlands and the London Green Belt in the UK. It was anticipated that these areas would provide the masses with

[1] Unless stated otherwise, the remaining part of this section is based on Dietvorst (1993a, 1987) and Beckers (1995 & 1983).

the opportunity to spend their newly obtained free time sensibly, and without unwanted damage to the countryside. In addition, these first concepts often served other purposes such as unemployment relief in the case of the "Amsterdamse Bos" or confinement of urban growth in London. Initiatives for these areas were often taken at the local level and projects were based on moral and conservation considerations rather than behavioural theories and models.

Following World War II, the demand for outdoor leisure facilities grew rapidly due to further individualisation, economic growth, urbanisation, population growth, increased leisure time and the development of mass mobility. Roadside picnicking became a very popular way of spending leisure time during the weekends. Compounded by concentration in time and space, the massive nature of outdoor recreation induced national planning agencies to put forward practical short-term solutions. Leisure, outdoor recreation in particular, was regarded as a spatial planning problem, and the planning concepts in the 1960s aimed at providing lasting availability of outdoor recreational opportunities in the vicinity of urban areas. Basically, the First (1960) and Second (1966) Reports on Physical Planning proposed a systematic and rational approach by establishing mono-functional 'large-scale elements for day-recreation' to facilitate recreational day-trips and to elevate the pressure on natural areas. In 1965, a new Ministry of Culture, Recreation and Social Work was established in the Netherlands that developed its own tools to advance outdoor recreation. Among the most important instruments was the Procedure for Project Development in Outdoor Recreation (1969) that advanced agencies for co-operation between local governments. The objective of these agencies was to develop and manage Designated Outdoor Recreation Areas (DORA's) at the regional level. Following the planning authorities, many commercial attractions (e.g., resorts and theme parks) were created at distant locations because of low land prices and the obvious psychological benefit of separation from the crowded urban cores (Stemerding, 1996).

In response to the rational and facilitative planning approaches in the 1960s, recreation and tourism research advanced from descriptive studies, through the identification of variables related to leisure behaviour, to various types of predictive models. The prediction stage required greater attention to the functional form of leisure behaviour models, including assumptions underlying different model specifications (Stynes & Peterson, 1986). In the 1960s and 1970s, when these supply-oriented planning approaches emerged, aggregate models such as gravity and entropy-maximising models and econometric demand models based on time series dominated the field. Consequently, early recreation and tourist models were concerned with interzonal interaction patterns rather than with individual choices (e.g., Baxter & Ewing, 1981).

During the 1970s, non-materialistic thinking influenced recreation and tourism. Conditions were created to enable people to spend their leisure time according to their own preferences. This new direction aimed to develop an integral view on leisure policies. However, at the dawn of the 1980s, the focus of policy makers changed due to economic recession and the emergence of environmental awareness. Social and economic aspects of recreation and tourism became more important (Inskeep, 1998), and planning for leisure developed a strong market orientation. At the national level, the government increasingly focused on development and growth by creating optimal conditions for market forces. This also included a strong international orientation aimed at increasing the European and global market share of the Netherlands as a holiday destination and at balancing domestic and foreign travel spending. Furthermore, due to the reappraisal of urban areas, the inner city became part of the leisure product in overall plans for recreation and tourism (Jansen-Verbeke, 1988). The 1986 Memorandum on the Policy on Holidaymakers marked the transitions to the new stage in recreation planning and policy. This Memorandum advanced, amongst others, diversification, accessibility, attainability, research, promotion, information and education. At the regional and local level, recreation and tourism policies were integrated and together recreation and tourism were expected to bring prosperity to economically weak areas. In these areas Tourism Recreation Overall Plans (TROPs) were introduced to co-ordinate and steer economic development. With regard to this stage of planning for leisure in the Netherlands, Beckers (1995, p. 96) concludes that spatial planning lost its political basis and its links with the ecological movement and that commercial tourism took the lead. The relationship between the state and the market seriously changed in the 1980s and Public-Private Partnerships become a popular concept to steer the expansion of high quality recreation and tourism facilities and services.

The research community at the time became aware of the fact that the popular gravity and time-series models from the 1960s lacked an underpinning consumer theory for their functional form and specification (Lieber & Fesenmaier, 1984; Peterson *et al*., 1983) and gradually these aggregate models were replaced by *disaggregate behavioural choice models* developed from random utility theory in psychology and economics (Timmermans & Golledge, 1990). The first Dutch publications introducing this novel approach to tourism and recreation were published in 1985 (Timmermans, 1985a, 1985b). Revealed preference and choice models and so-called decompositional multiattribute preference models (or conjoint analysis) were among the most popular disaggregate models. Revealed choice models explain observed choice patterns in terms of an underlying utility function. Conjoint analysis assumes individuals to cognitively integrate their evaluations of a choice alternative's attributes to obtain some overall utility for that alternative. The choice alternative with the highest utility is assumed to be selected (Timmermans & Golledge, 1990). Compared to revealed preference and choice models, conjoint models have the advantage that they are based on carefully designed experiments. This approach is also referred to as "stated" preference and choice analysis and it is very useful for the evaluation of policy measures because it is possible to obtain data on choice alternatives presently not available in real-world markets (Louviere & Timmermans, 1990a, 1990b).

1.2 Modelling challenges at the dawn of the 21st century

In line with the transition to a demand- and market-orientation on recreation and tourism, the 1986 Memorandum and the subsequent policy documents (a.o., "Enterprise in Tourism" (1990), the 1991 "Policy Report on Outdoor Recreation" and the 1993 "Structure Scheme on Green Areas") catered to the trends that have characterised Dutch leisure consumption since the mid-1980s. Apart from the ongoing growth of leisure activities and expenditures, one of the most salient trends has been that modern consumers increasingly wish to put together their leisure arrangements in an efficient way and desire custom-made arrangements based on pre-existing choice opportunities. "It seemed that with increased leisure time people were increasingly moving away from the somewhat standardised package holiday and seeking out a wider variety of forms of leisure activity, including independent travel" (Urry, 1990, p. 50).

More generally, there has been an increasing desire for variety and diversification[2] in (outdoor) leisure activities. As an example of variety seeking, for instance, (fun) shopping has become more significant to recreation and tourism, both in terms of spending and as an incentive for day tripping and the selection of holiday destinations. The need for variety has contributed to the increased importance of short breaks, of long-distance travelling, and of holidays and day-trips outside the traditional holiday seasons. In this context, at least three aspects of variety in our (post)modern society have affected the leisure related-use of space

[2] "Consistent with the marketing literature (..), *diversification* is variety that takes place within a set of choice alternatives that is within a well-defined and specific time period (i.e. a day visit to a park), while *variety seeking* occurs over longer periods of time (i.e. between different visits to a park)" (Kemperman, 2000, p. 4).

and time (Dietvorst, 1993b): (1) the variety in the types of households; (2) the variety in leisure behaviour within households; and (3) the variety in leisure needs and preferences of each individual. According to Dietvorst (1993b), this has resulted in an increase in the number of market segments and the need for entrepreneurs and policy makers, to take into account the desires and expectations of many different users' groups.

The increased complexity of leisure behaviour and demand for leisure goods and services that has been evident since the mid-1980s has led to a new orientation in recreation and tourism policies and research. In the research community, the first international exploratory studies on the relationships between various facets of leisure trips were published during the second half of the 1990s (e.g., King & Woodside, 2001; Jeng & Fesenmaier, 1997; Woodside & MacDonald, 1994). It is now realised that leisure behaviour results from complex interactions between individual and household preferences, styles and constraints, the institutional context (e.g., working and school holiday regimes), and the availability of (leisure) facilities and services. In other words, leisure behaviour is inextricably related to the goals individuals' and households' aim for in daily-life. Moreover, it is now realised that the travel decision itself is not singular but rather comprised of a number of interrelated sub-decisions including the members of the travel party, timing, mode of transport, accommodation, destination choice and so on. Finally, it now believed that subsequent trips affect each other (Dellaert *et al*., 1998; Lindh *et al*., 1995). As a result, research is now evolving towards the analysis of activity patterns - clusters of interrelated (leisure) activities that are pursued within a certain period of time and under various constraints. The so-called *activity-based approaches to travel forecasting*, for instance, aim to predict which activities are conducted where, when, for how long, with whom and with which transport mode (see for an overview: Timmermans, 2000; Ettema & Timmermans, 1997; Ettema, 1996)[3].

In the light of the recent focus in leisure planning and research, the conventional disaggregate utility-based models do not longer suffice because they focus on single trips and capture only part of the travel decision. Consequently,

[3] Patterns of leisure behaviour have also been explored using sociological and psychological concepts such as life styles, tourist, recreation and leisure styles, attitudes, experiences and/or motives (e.g., Te Kloeze, 1990; Glyptis, 1989; Cohen, 1979). With regard to *forecasting* future behavioural patterns, however, several studies have shown the predictive power of these concepts and segmentations to be rather disappointing in the context of recreation and tourism in the Netherlands (e.g.,. Van Middelkoop, 1996; Van Keken *et al*., 1995). This thesis therefore concentrates on the more 'traditional' socio-stuctural conditions within which tourist decision-making processes take place.

they ignore the interdependencies between subsequent leisure trips and between these trips and other aspects of daily life. Also, they are unable to describe time-related choices satisfactorily. Some initial impetus has been given to solve these problems *within utility-based disaggregate models* (e.g., Kemperman *et al*., 2000; Dellaert *et al*., 1997; Fisher *et al*., 1990; Timmermans & Borgers, 1989; Fesenmaier, 1985), but again these attempts only comprise part(s) of the complex leisure decision-making process.

The activity-based approaches to travel forecasting have also led to the development and application of new models that are not based on utility-maximising theory. *Constraint-based models*, for instance, are based on the idea that people are not free to choose the activities they desire because they have to take into account restrictions such as limited time and money budgets or coupling constraints. These approaches descend from Hägerstrand's time geography (Hägerstrand, 1970), and since about 1980, a growing number of papers has appeared dealing with the interconnected issues of non-participation and constraints in leisure activities (Jackson, 1988).

Alternatively, adherents of *qualitative* approaches withdraw from the assumption of utility-based models that decision-makers have complete information of all alternatives and that they are able to arrive at optimal solutions. Instead, based on modern psychological and physiological theory, qualitative approaches assume individuals to use heuristic decision-making rules that represent decisions that have worked out satisfactorily under similar conditions in the past. These decision rules reflect both desires and constraints of the decision-maker. Qualitative models offer more flexibility in modelling choice behaviour because, in contrast to the quantitative utility-based approaches, they do not impose rigorous assumptions on the data. Like the traditional quantitative models, however, they also focus only on one, or at best a couple of parts of complex leisure decisions.

In conclusion, therefore, it can be stated that, despite the tremendous progress in recent years, both traditional utility-based approaches and approaches based on modern psychological and physiological theory still lack the sophistication required to gain insight into individual leisure trip patterns. At the same time, the need for more comprehensive models of outdoor leisure choice behaviour and demand is increased due to the still growing market orientation of policy makers and industry alike. To really understand these complex processes, it might be better to develop much richer models, composed of more variables, and representing more complex relationships among those variables. In particular, this applies to:

(1) the number and type of subdecisions (facets) of the leisure choice process and the relationship between the facets of one leisure trip;
(2) the type of leisure trips included in the trip pattern, in particular the relationship between day-trips and overnight holidays; and
(3) the relationship between (the facets of) leisure trips of one individual (i.e., influences of trips that are separated in time on each other).

This thesis therefore aims to fulfill this need for more comprehensive models by developing a *simulation model of outdoor leisure behaviour*. Simulation allows the modelling of a system in almost limitless detail (Levine & Lodwick, 1983b), because it allows analysts to construct approximations of complex real-world situations by synthesising from relatively simple parts (Shannon, 1975). Given the present interest in the complex leisure decision-making process, *micro*simulation in particular appears to be a promising line of development, because microsimulation models are directly concerned with the behaviour of microunits (e.g., individuals, households or firms) in different phases of the choice process (Merz, 1991). In other words, microsimulation offers the opportunity to build more comprehensive models of leisure choices by combining behavioural hypotheses and relationships with regard to the facets that comprise this decision. In this respect, both the traditional and the more recent approaches can be used to describe the identified choice facets. Aggregate effects, then, are obtained by combining individual simulations.

1.3 Research objective and questions

In order to develop a tool that will provide a fundamental understanding of the impact of policy measures and autonomous developments on outdoor leisure trip patterns, this thesis aims to *build a comprehensive microsimulation model of outdoor leisure behaviour that is based on a representation of the underlying decision-making processes and that will assist planners in evaluating alternative scenarios and planning options.* This comprehensive representation should emphasise (1) the interrelations between individual and household preferences and constraints; (2) the various facets of outdoor leisure trips; (3) the effects of leisure trips that are separated in time on each other; and (4) the influence of the physical and institutional leisure environment. Leisure trips are studied over a longer period of time and the "trip patterns" under consideration are defined as the set of interrelated day- and overnight leisure trips that are pursued by an individual within that period of time. These outdoor leisure trips are referred to as "tourist trips", because tourism definitions are traditionally rooted in notions of economics, time use and distance travelled. It is noted, however, that some of the day-trips studied would be better characterised as “recreational” trips.

In short, the model under development is a microsimulation model of tourist trip patterns that is based on the underlying decision-making processes of tourists, and that is more comprehensive than existing models in terms of the number and type of facets and trips that are included. In order to develop such a comprehensive microsimulation model by synthesising from the (simple) parts of the tourist decision-making process, three issues are important (partially based on Shannon (1975)). First, a comprehensive conceptual representation of the tourist decision-making process should be developed, including the types of trips, the facets of these trips, and the interactions between these facets and trips that are relevant to the tourist trip patterns under consideration. Next, these trips, their facets and the interactions should be correctly related to the conditions that structure the tourist decision-making process. Finally, these relationships should be synthesised to build a comprehensive model of the tourist decision-making process that allows planners and policy makers to evaluate alternative scenarios and policy options in terms of the entire tourist trip pattern. More specifically, the following research questions should be answered in this thesis:

Towards the conceptual building blocks of the simulation model:

(1) How are tourists’ decisions made? What facets (subdecisions) of tourist trips (frequency, duration, destination, travel party, accommodation and so on) are relevant, and how do these facets and trips relate to each other?

Towards the empirical building blocks of the simulation model:

(2) How can each facet of the tourist decision-making process be modelled best? And how can the relationships between these facets be included in these empirical models?

Towards the correct structure of the simulation model:

(3) How can the empirical models be combined to simulate the tourist decision-making process for tourist trip patterns?

1.4 Outline of the thesis

The proposed microsimulation model of tourist trip patterns is part of a decision support system, referred to as the *MERLIN*-system, which will assist planners in evaluating alternative scenarios and planning options. *MERLIN* bears reference to the once and future magician, who's figure combines the roles of a wise man, seer, prophet and shaman and who has walked the stage of literature and tradition since Celtic times (Stewart & Matthews, 1995). In contrast to the historical figure whose actions and purpose were seldom wholly revealed, the *MERLIN*-system will be developed systematically and documented properly like an academic thesis should.

In order to achieve this objective, this thesis is organised as follows. Chapter 2 reviews the contributions of the tourism literature to our understanding of tourist decision-making today. Based on this discussion and in reply to research question 1, a conceptual framework for day- and overnight tourist trips is presented. Next, chapter 3 reviews the various modelling approaches that have been applied to recreation and tourism in the past. Based on this review, research question 2 is answered. Subsequently, chapter 4 discusses the principles of simulation, including an overview of existing (micro)simulation models in recreation and tourism research. This chapter thus offers a review of existing solutions to building more comprehensive models (research question 3). Concluding the chapters on the conceptual development of the *MERLIN*-system, chapter 5, finally, formalises the modelling problem and outlines the architecture of the model. This includes a detailed description of the components that embody the conceptual representation of the tourist decision-making process, and the components that control the simulation process.

Having defined the *MERLIN*-system, chapter 6 discusses the data that were collected to calibrate the empirical building blocks of the system. Using these data, chapters 7 through 11 inclusive discuss the models that describe the various facets of the tourist decision-making process. Chapter 12, then, concludes the development of the system by discussing the final empirical adjustments and by validating the performance of the entire system in terms of its ability to reproduce the original data at both the aggregate and the individual level. Subsequently, this chapter demonstrates how *MERLIN* can be used to simulate scenarios for the future and assess the effects of possible policy measures and general developments in society on tourist trip patterns. Chapter 13, finally, concludes this thesis by summarising the major research findings and identifying potentially promising areas of future research.

2 The Tourist Decision-Making Process

Typically, (quantitative) models of tourists' choices are based on conceptual representations of the tourist decision-making process. This chapter first discusses a general model of consumer choice behaviour and then reviews the contributions of the tourism literature to our understanding of tourist decision-making. Based on this discussion, the conceptual framework for the *MERLIN*-system is presented.

2.1 Introduction

This thesis aims at developing a more comprehensive model of tourist choice behaviour that is based on the underlying decision-making processes of tourists. In order to attain this goal, we need to understand how tourist decisions can be conceptualised (Mitra & Lankford, 1999; Smith, 1995; Mansfeld, 1992). This is the topic of consumer behaviour, an interdisciplinary science that studies the decision-making activities of individuals in their consumption roles (Schiffman & Lazar Kanuk, 1991). This chapter therefore first discusses a general model of consumer decision-making. The subsequent section reviews the specific contributions of tourism research to our understanding of tourist decision-making. First, it will discuss conceptualisations of tourist destination choices. Secondly, based on more recent publications, it develops a more comprehensive view on tourist decision-making and argues that the tourist decision is not a singular decision, but rather is comprised of a set of interrelated sub-decisions that evolve in time. Furthermore, it will be argued that the tourist choices can be viewed as the product of complex interaction processes between constraints, opportunities and motivations. Building on this more comprehensive conceptualisation of tourists' decisions, the final section presents the conceptual framework underlying the *MERLIN*-system.

2.2 A general model of consumer decision-making

A widely accepted model of consumer decision-making for *high-involvement, non-routinised* purchases comprises five-stages: (1) problem recognition and formulation; (2) information search; (3) evaluation; (4) purchase or implementation; and (5) post-purchase evaluation (Chen, 1997; Crompton, 1992). These stages represent the *process* and the *output* components of consumer decision-making (see Figure 2.1). In addition to process and output, an *input* level is identified, which draws upon external influences that serve as sources of information about a particular product and affect the consumer's product-related values, attitudes, and behaviour (Schiffman & Lazar Kanuk, 1991).

Two external influences are distinguished, including marketing-mix activities (e.g., product, promotion, price, place, personnel and planning), and social cultural influences. In addition to Schiffman and Lazar Kanuk's (1991) external influences, Figure 2.1 also includes government policies and consumer characteristics at the input level. Government policies are included because they affect supply conditions and personal and household constraints arising from, for instance, working and school holiday regimes. Consumer characteristics, although strictly not external, are included because they influence the evaluation of choice alternatives, e.g. a consumer will not consider using a car when (s)he does not have a driver's license.

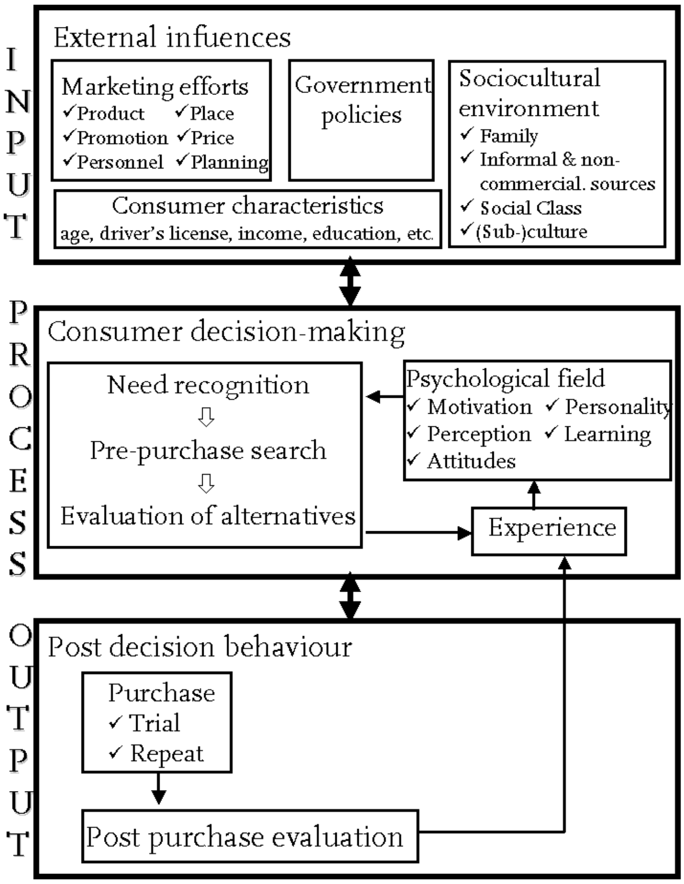


Figure 2.1 A general model of consumer decision-making (based on Schiffman and Lazar Kanuk (1991, p.557-572))

The *process* component of the general model of consumer decision-making is concerned with the actual consumer decision-making process. First, a consumer is faced with a problem (actual state recognition) or the desire to buy a certain product is triggered (desired state recognition; e.g., by promotional activities). Next, the consumer engages in the stage of pre-purchase search using both internal sources (past experiences) and external sources (marketing, non-commercial and social communications) to collect information. The information search process includes the establishment of a set of alternatives to be considered (the evoked set) and their attributes, as well as the set of criteria against which these alternatives will be evaluated. In the third stage of the decision-making process, the consumer will assess the alternatives in the evoked set using the established criteria and certain consumer decision rules. These decision rules are also referred to as heuristics, decisions strategies, or consumer information-processing strategies[4]. Since the actual choice is made at this third stage of the process component, the quantitative choice models that will be discussed in the next chapter often represent this part of the decision-making process.

The final stage in consumer decision-making refers to the *output* or results of the process. This component has two substages - the actual purchase of a product or service (or behavioural (external) response), and the postpurchase evaluation (or internal response (Hansen, 1976)) - both aiming at increasing the consumer's satisfaction with his or her purchase decision. Purchase behaviour can be further divided into trial purchases and repeat purchases, or, in terms of tourism destination choices, “first time and return visits”. Both trial and repeat purchases feed the consumer's experiences following a process of postpurchase evaluation. These experiences will influence the consumer's future decision-making processes as the consumer develops and adjusts his motivations, perceptions, personality and attitudes.

It should be emphasised that the presented general model of consumer decision-making is a simplified representation of decision-making processes under high-involvement, non-routinised conditions. In particular, the described choice sets structure is most likely to be useful in high-involvement destination choices that will be discussed in the next subsection. In situations where the (potential) consumer is less involved with his or her decision (e.g. short holidays or repeat visits), the selection decision need not include all components previously described (Crompton, 1992), or the order of the components may be slightly different. On the other hand, in choice situations that comprise many high-involvement choice

[4] Basically, there are two types of consumer decision rules. In following a *compensatory decision rule*, consumers simply add the weighted attribute evaluations (utilities) for each product in the evoked set. Consequently, this decision rule allows for compensation of positive and negative evaluations of a product's attributes. In contrast, *non-compensatory decision rules* do not allow for the balancing of positively and negatively evaluated product attributes. It has been found in marketing science, management science and psychology that, as the number of choice alternatives and/or the number of attributes increases, individuals make greater use of non-compensatory decisions strategies (Olshavsky, 1979; Payne, 1976; in: Timmermans & Van der Heijden, 1987).

facets (destination, travel party, transport, etc.), the choice process is likely to be more complex, including, for instance, simultaneous reflections on different destination alternatives and different modes of transport.

2.3 Tourist decision-making processes

Early models of recreation and tourism behaviour rarely paid attention to the tourist decision-making process. It was not until the 1970s that the concepts and models of consumer behaviour were applied to tourist decision-making. Initially, research focused on (the structure of) tourist destination choices. Recently, however, the first exploratory studies on the relationships between various facets of tourist trips have been published. It is now generally recognised that the travel decision-making process comprises a number of interrelated trip decisions or *facets*, and is better described as a *trip profile*, including decisions regarding the members of the travel party, timing, transport mode, accommodation, destination choice and so on. In addition, the complex interaction between preferences, opportunities and constraints in the tourist decision-making process has recently drawn the attention of the research community. This section first reviews research on destination choices. Next, a more comprehensive view on tourists' travel decisions is discussed. Based on this discussion, section 2.4 presents a conceptual framework for day- and overnight tourist trips that provides the conceptual basis for the *MERCIN*-system

2.3.1 Tourist destination choices

Since the 1970s, a number of destination choice models have been advanced that tried to explain tourist destination choice behaviour. The majority of these models fairly resemble the model of consumer decision-making that was discussed in the previous section. In this case, the "consumer product" is a tourist destination and the "purchase" represents the visit to that destination. The following discussion focuses on the specific contributions of various authors to our understanding of tourist destination choices.

Consumer choice processes for non-routinised, high-involvement purchases are reported to be phased (Crompton, 1992, p. 421). Um and Crompton (1990) conceptualise the tourist destination choice process as having two phases. First, the potential tourist has to decide whether or not to have a holiday. This can be interpreted as the *participation choice*. Conditional upon this decision, the second phase is concerned with the question where to go, the actual *destination choice*. This phasing corresponds to recent advances in recreation demand that assume participants to first decide on the number of trips to take per time unit (e.g., a season), while the second stage of the decision model is concerned with the allocation of these trips across substitute sites (Siderelis & Moore, 1998; Parsons & Kealy, 1995; Hausman *et al*., 1995).

Mansfeld (1992) argues that an analysis of the *motivational stage* that sets off the whole decision-making process can reveal the way in which people set goals for their destination choice and how these goals are then reflected in both choice and travel behaviour (Mansfeld, 1992, p. 401). Basically, there are two types of motivational factors, including *push factors* that make people want to travel, and *pull factors* that determine where people want to travel given their initial desire to travel. Many have tried to identify these factors. Iso-Ahola and co-authors, for instance, argue that two motivational forces simultaneously influence the individual's leisure behaviour, including escaping (inter)personal environments and seeking (inter)personal rewards (Mannell & Iso-Ahola, 1987; Iso-Ahola, 1984, 1982). However, Mansfeld (1992) and Witt and Wright (1993) conclude that the various lists of travel motives cannot satisfactory predict tourist destination choice behaviour because travel decisions are often 'multi-motive' situations, and because other factors influence the decision both directly and indirectly (e.g. a person's budget, beliefs about the weather, etc.).

Elaborating the choice sets structure of destination choices, Woodside and Lysonski's (1989) general model of traveller destination choice focuses on awareness, preference and choice of competing destinations. More specifically, the model distinguishes four mental categories, including (1) the *consideration set*, corresponding to the evoked set that was discussed previously; (2) the *inert set*, for which the consumer has neither a positive nor a negative evaluation; (3) the *unavailable and aware set* of destinations that are difficult to go to; and (4) the *inept or reject set* of destinations that the consumer has rejected because of negative previous experiences or negative evaluations from other information sources (Woodside & Lysonski, 1989).

Although many models refer to the influence of previous experiences, few studies have focussed on longer periods of time. Oppermann (1998) found that respondents often did not continue to travel to the most distant destination zone once they had done so. Also, many respondents often revisited destinations within the same destination zone, suggesting that they were unwilling to explore and 'risk' other destination zones (Oppermann, 1998, p. 328). Individual travel careers are most likely to be explained by traveller characteristics and life cycle changes (Oppermann, 1998, 1995; Lawson, 1991) or cohort effects (Oppermann, 1998).

2.3.2 A more comprehensive discussion of tourists' travel decisions

Today, it is recognised that travel decisions comprise more than just destination choices. In the field of tourism, Woodside and MacDonald (1994) were one of the first to describe this new perspective on tourist choice processes. Their "Systems Model" includes the input, process and the output components that have been described previously. In addition, it also details the process component by identifying eight *leisure travel choice subsystems,* including destinations, accommodations, activities, visiting attractions, travel modes/routes, eating options, destinations areas and routes and self-gifts and other durable purchases (see Figure 2.2). The authors used their model to analyse long-interview data, and concluded that "compared to linear models (..) the proposed general systems framework is a rich, contextual foundation for deeper knowledge and insights of why and how specific traveller-related decisions and behaviours occur."

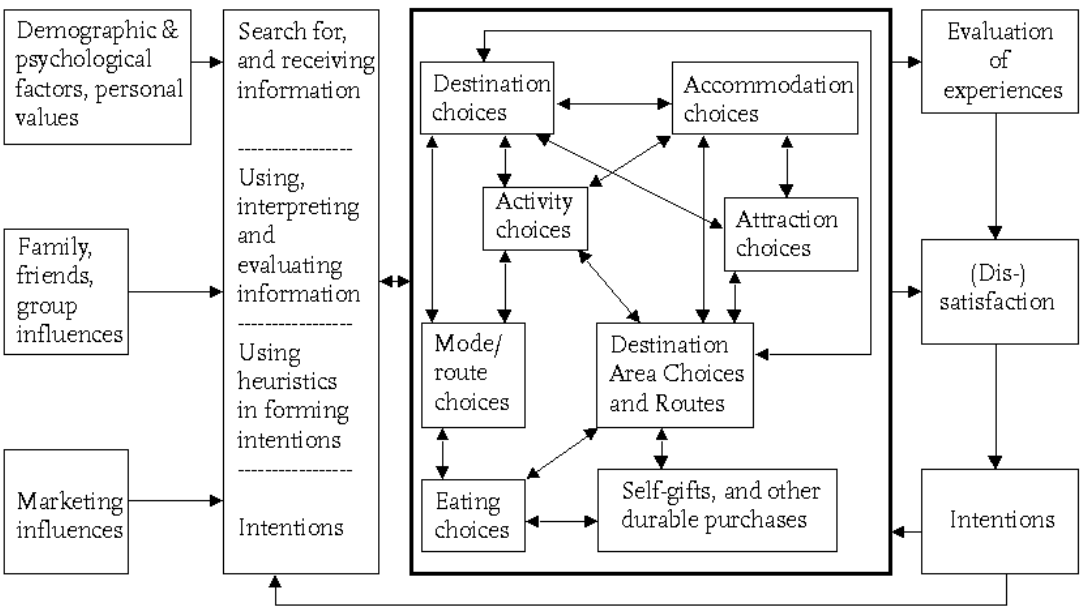


Figure 2.2 Woodside and MacDonald's (1994, p. 33) "General systems framework of customer choice decisions of tourism services"

Later studies focussing on tourist choice processes referred to Woodside and MacDonald's subsystems as trip (sub)decisions, dimensions, components or facets, and the whole travel decision is now often described in terms of a trip *"profile"*. Also, the number of identified inter- and intra-trip relationships has gradually been expanded. First, trip profiles have been extended so as to also include the choice of the travel group (a.o. Dellaert *et al*., 1998a; Jeng & Fesenmaier, 1997), budget considerations (Jeng & Fesenmaier, 1997), the length of the trip (Dellaert *et al*., 1998a; Jeng & Fesenmaier, 1997; Dadgostar & Isotalo, 1992) and trip frequencies and timing (Bargeman *et al*., 1999).

Secondly, several studies have focussed on how tourists temporally separate their choices of facets. Dellaert *et al.* (1998a), for instance, found that destination and travel companionship choices are typically made 6 to 7 months prior to the actual travel date, accommodation and trip duration are selected approximately 5½ month before departure, the exact date of departure is set 5 months ahead, whereas bookings are not made until 3 to 4 months before the actual travelling activity. It should be noted, however, that according to this study, these planning horizons do not differ significantly from each other, because of the rather large standard deviations of these averages across the sample of trips. Wijker (1998), on the other hand, argues that the tourist-consumer first determines the duration of the activity. Given this decision, (s)he considers the timing (season), the choice of the destination (depending on the activities that can be pursued in the area), and finally, the accommodation.

Another study conducted by Jeng and Fesenmaier (1997) among Midwest USA residents showed location of overnight stay, primary destination, date, timing and duration of the trip, and travel partners to be at the core of the travel decision. These decisions are taken long before departure. In contrast, food and rest stops and shopping proved to be peripheral elements in this study while these trip decisions are often left for *"en route"*-decisions. Finally, this study indicated trip decisions that sit between the primary and peripheral spheres to be most fuzzy. Perhaps these decisions are important to differentiate between different trip types (such as shorter vs. longer trips)[5].

Third and finally, the interrelations between tourist trips that are separated in time have attracted the attention of the research community. First, a distinction between principal and additional holidays has been used to explain the relative importance of certain facets in the tourist decision-making process. Based on a survey, Dirven *et al*. (1998) concluded that for most people (85%), the most important holiday is the longest holiday. A minority, however, considered the most distant (10%) or the most expensive (5%) holiday to be the main holiday. Secondly, the influence of (the facets of) subsequent tourist trips on each other has been explored. The trip purpose and travel mode, for instance, were found to influence the interval between two subsequent trips (Lindh *et al*., 1995).

In addition to the broadening of the *content* of the tourist choice process, our understanding and appreciation of the complexity of the *decision-making process*

[5] Trip duration was not included in Jeng and Fesenmaier's study because they had asked their respondent's to select a destination in the midwest USA for a short (2-4 days) summer holiday.

itself has also increased over the past decades. In particular, the interaction between preferences and constraints in the decision-making process has received more attention. Advocates of *choice-based theories* consider observed behaviour as an expression of people's preferences. Descending from Hägerstrand's time geography (Hägerstrand, 1970), however, *constraints theories* argue that people are not free to choose the activities they prefer because they are restricted by constraints. In recreation and tourism research, constraint-related phenomena have often been examined in connection to the issue of nonparticipation. The type of constraints that have been investigated include work commitments, family commitments, lack of awareness and/or absence of opportunities, poor facilities, lack of money, lack and cost of transport, cost of equipment, lack of interest, physical disabilities, lack of time, lack of partners, admission fees, shyness, safety concerns, crowding and pollution of sites, and geographical constraints such as the ability and distance to leisure services (Jackson, 1988). The traditional view was that "first a leisure preference exists, then a barrier intervenes and results in non-participation or, if no barrier intervenes, the individual will participate" (Crawford & Godbey, 1987). Later research toned down this view by suggesting that (leisure) participation is dependent not on the absence of constraints but on negotiation through them. Leisure participation is now viewed as the product of a balance between constraints and motivations (Jackson *et al*., 1993). In this view, constraints may act differently on tourist preferences depending on:

(1) their source: *internal* (also referred to as personal or capacity constraints) or *external* (also referred to as social, coupling and/or authority constraints) (Dellaert *et al*., 1998a; Jackson, 1988; Hägerstrand, 1970);
(2) their intensity (e.g., blocking or inhibiting);
(3) the stage of the decision-making process during which they are negotiated: "intra- and interpersonal" (Crawford & Godbey, 1987) or "antecedent" constraints (Jackson, 1990) influence individual preferences because the anticipation of insurmountable interpersonal or structural constraints may suppress the desire for participation, whereas "intervening" or "structural" constraints (Crawford & Godbey, 1987) step in between preferences and (the) participation (choice) only).

To conclude, it should be noted that, in addition to different types of constraints, different coping (or negotiation) strategies have been reported. Kay and Jackson (1991), for instance, identified three groups of responses:

(1) people who do not participate in their desired activity (*reactive response*);
(2) people who, despite experiencing a constraint, do not reduce or otherwise change their participation at all (*successful proactive responses*); and

(3) people who participate but in an altered manner (*partly successful proactive responses*).

In contrast to these *behavioural* strategies, *cognitive* strategies (such as the working of antecedent constraints), aim to reduce the psychic (dis)comfort of the decision maker (Jackson *et al*., 1993).

Building on this more comprehensive conceptualisation of tourists' travelling decisions, the final section of this chapter presents the conceptual framework underlying the *MERLIN*-system.

2.4 The conceptual framework underlying the *MERLIN*-system

This section presents the conceptual basis for the *MERLIN*-system. The framework is presented in Figure 2.3 and represents the decision-making context for tourist trip patterns. In this context, a *tourist trip pattern* is defined as an interrelated set of tourist trips, each of which is characterised by a *trip profile* (i.e., a description of a trip in terms of various facets), that is pursued by an individual during a particular period of time. In case of the *MERLIN*-system, the period under consideration is one year, and the tourist trips include both day- and overnight trips.

Like all economic activities, the decision-making context for tourist trip patterns can be organised into supply and demand side factors (Holecek, 1993). On the *supply-side* of tourism, governments and public authorities, private enterprises and the non-profit sector together determine the available tourist-recreation facilities and services (Van Lier, 1993). This broad and complex set of opportunities is referred to as the *Tourist-Recreation Product* (TRP). Depending on the function within the TRP, three elements can be distinguished (Jansen-Verbeke, 1988); without these elements, one cannot speak of a true "product". First, "*primary elements*" are those attributes and facilities that attract tourists and visitors. Examples include amusement parks, museums, cultural and sports facilities, events and festivities and extraordinary landscapes. Secondly, the TRP is embedded in a physical and organisational framework that supports the primary elements, but that does not exert an autonomous attraction on visitors itself. These so-called "*secondary elements*" include, for instance, the hotel and catering industry, shopping facilities and markets. Finally, "*additional elements*" complement primary and secondary elements by facilitating local, regional or national tourist industries. These facilitating elements include transport systems that contribute to the accessibility of the product and information and promotional elements that enhance the tourists' awareness and knowledge of the available opportunities

(Jansen-Verbeke, 1988). When the elements within the TRP are also spatially related, that is when tourists combine the elements in their movements, the TRP can also be described as a system, or a *Tourist Recreation Complex* (Dietvorst, 1995, 1993b, 1989). A tourist-recreation complex is a spatially differentiated whole and it may have different spatial scales (Dietvorst, 1995). Examples of complexes at different scales include areas (e.g., Spaarnwoude), cities (e.g., Amsterdam), regions (e.g., the lake area in the southwest of Friesland) or even whole countries. Within these complexes the positions of the product's elements can be described as subordinate, co-ordinate, complementary, reinforcing and so on (Dietvorst, 1995).

The *demand-side* of tourism, on the other hand, is concerned with the tourist-consumer of the TRP. From the perception of the individual consumer, both constraints and preferences determine how a (potential) tourist will allocate the available resources to leisure activities. Many constraints are determined by long term decisions of the individual and his or her household. These decisions, including, for instance, education and work, civil status and the acquisition of means of transport and tourist commodities, often take a long time to change. Also, they are often related to life style considerations of the individual and his or her household. In the conceptual framework for tourist trip patterns presented in Figure 2.3, these decisions are represented by the personal and household characteristics and the mandatory activities. In terms of constraints they can be further specified as (see o.a. Arentze & Timmermans, 2000c; Hägerstrand, 1970):

(1) *capability* constraints due to physical incapacities and delicate health or the availability of commodities and skills (e.g., a car or a driver's licence);
(2) *coupling* constraints due to household composition and activities of other household members;
(3) *financial* constraints will limit the expenditures on leisure activities;
(4) *time-budgets* will constraint the number of days available for (out-of-home) leisure activities; and
(5) *space constraints*, e.g., pre-scheduling arrangements such as seasonal reservations for a caravan.

Other constraints in the decision-making process arise from sources outside the tourist, including the supply-side of the market. They include:

(6) *authority* constraints, such as availability and accessibility of facilities and services or school and national holiday seasons;
(7) *logical* constraints that limit the availability of alternatives within choice sets (e.g. a person without a driver's licence cannot decide to go on a car-based holiday without other travel partners);
(8) *spatio-temporal* constraints, e.g., people in Europe cannot engage in a 2-days

short break to Australia by car; and

(9) *situational* constraints, such as the weather, will affect the decision-making process in the implementation phase.

We will refer to the first set of constraints (1-5) as "*personal and household constraints*" and to the second set (6-9) as "*system and institutional constraints*".

Constraints can also influence preferences because people may adjust their preferences to anticipate constraints and anticipate dissatisfaction. These antecedent constraints (Jackson, 1990) are indicated in Figure 2.3 by the relationships between the personal and household preferences and the TRP on the one hand and personal and household constraints on the other.

In the tourist decision-making process, supply and demand side conditions (i.e., personal and household and system and institutional constraints and opportunities) are negotiated by the (potential) tourist based on his/her personal and household preferences with regard to leisure and to other facets of (daily) life. In short, the selection of the annual set of tourist trips is interpreted as an allocation problem in which the decision to pursue tourist trips is traded-off against other time and money consuming products. Five stages can be distinguished in this process, and at each stage the (potential) tourist will trade off his or her constraints and opportunities. First, participation in leisure activities in general is contingent on the time and money available after fulfilment of the essential subsistence and maintenance activities such as working, going to school and satisfying personal or household physiological and biological needs (Bhat & Koppelman, 1993). Second, given the resources available for leisure activities, the individual may distribute these time and money budgets among in-home and out-of-home leisure activities. Third, the resources allocated to out-of-home leisure activities may be spend on tourist trips and on non-tourist trips, such as visiting relatives and friends and short recreational outings. Fourth, given the resources available for tourist trips, the tourist needs to consider the preferred allocation of these budgets among day-trips and overnight holidays. Fifth and finally, the tourist will further decide on the profiles of the desired day- and overnight tourist trips (indicated by the layered facets x, y and z in Figure 2.3).

Given these stages, the position of the *MERLIN*-system within this framework, including some definitions, can be explicated. First, the *MERLIN*-system only distinguishes between resources allocated to tourist trips and those allocated to non-tourist activities and trips. As a consequence, the first three stages of the decision-making process as described above are not detailed by the system.

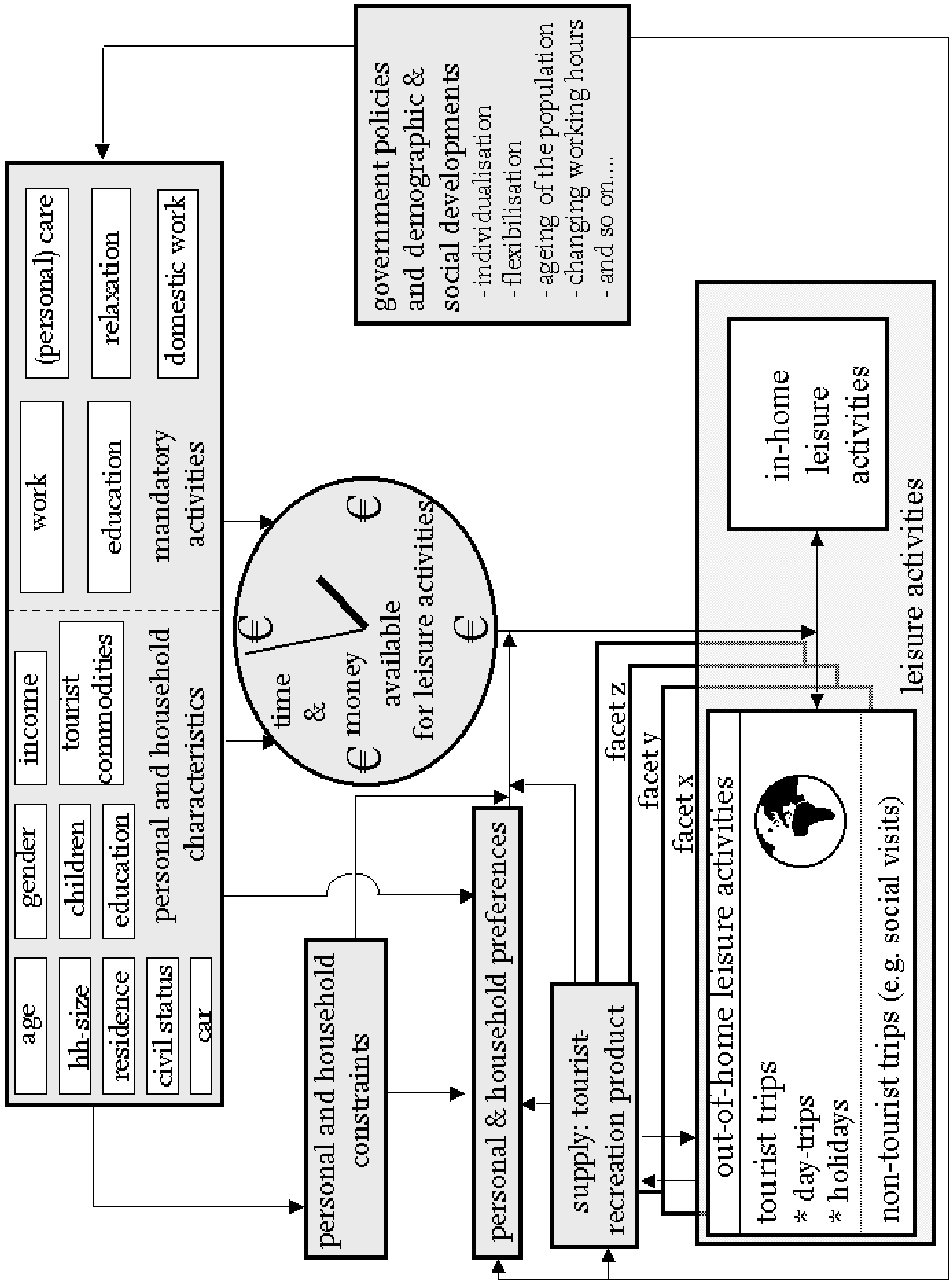


Figure 2.3 A conceptual framework for the generation of tourist trip patterns

Second, the MERCIN-system only comprises the facets of the travel decision that are (often) taken long before departure and/or that are relevant to policy makers and planners at the national level. Based on the review of the literature (section 2.3.2), these choice facets typically include the travel party, the timing of the trip, the destination, the accommodation, the transport mode, the date of departure and the expenditures. The *MERLIN*-system assumes a sequential decision-making process with regard to both the contemplation of main and additional tourist trips (where more important trips are considered first) and the choice(s) of facets of these trips. The order of the trips and facets was determined based on long-interviews conducted in 1997 that were inspired on the results of Jeng and Fesenmaier (1997). Later the order of facets was (at least partially) endorsed by publications by Dellaert *et al*. (1998a) and Wijker (1998). Chapter 5 will discusses this process in more detail and formalise it properly.

Third, the *MERLIN*-system defines *day-trips* as outdoor activities from the residential location for recreational purposes, involving at least four successive hours without spending a night. Trips starting at a holiday address are thus excluded, as are visits to family and friends. This corresponds to the definition of day-trips applied by the Dutch National Bureau for Tourism (NBT, 1989) and deviates from the more frequently applied definition introduced by the Dutch Central Bureau for Statistics (CBS). In its quintannual survey, the CBS denotes a two hours trip as a day-trip (CBS, 1992; CBS, 1997). The four-hour definition used in this thesis excludes frequent and periodical activities such as regular active sporting.

Holidays, on the other hand, involve overnight activities for recreational purposes. Again visits to family and friends are excluded, as are business trips (journeys on work-related grounds). However, the simulation model could easily be extended to also include trips with non-tourist motives. This definition is in accordance with the Continuous Vacation Research (CVO), an extensive consumer panel managed by the Dutch Research Institute for Recreation and Tourism (NRIT) and the Dutch Bureau for Tourism (NBT) (Van der Most, 1996; CBS, 1996).

Finally, in addition to the above definitions, the following assumptions are made with regard to the scope of the *MERLIN*-system. First, the time horizon of the tourist trips that are considered is *one year* and the simulation unit is *the individual*. Tourist trips of possible household members are not simulated, but the presence of other household members (if any) is assumed to have an effect on the individual's tourist choices.

Given this representation of the decision-making process for tourist trip patterns, the *MERLIN*-system will simulate a choice process in which an individual allocates a certain amount of (time and money) resources to day- and overnight tourist trips with certain characteristics (choice facets) during a one-year period given a set of personal and household and system and institutional constraints. This process constitutes the core of the MERCIN-system. Simulation systems like the MERCIN-system are, however, typically designed to assess the impact of particular changes on choice patterns, in the present case tourist trip patterns. To conclude the discussion of the conceptual framework, therefore, the two most important sources of change will be discussed. First, policy makers and analysts are interested in the effects of changes on the supply-side (i.e. the TRP), because suppliers in tourism aim to innovate their products in order to (seek to) anticipate or direct trends in leisure behaviour. In terms of the conceptual framework: changes on the supply side change the system and institutional conditions for the behaviour under considerations. Second, general developments in society will affect tourist trip patterns because they change both personal and household and system and institutional conditions. First and foremost, government policies and demographic developments will affect the personal and household characteristics of the population, including the resources allocated to their mandatory activities. This way, these changes affect the personal and household constraints, and, due to antecedent constraints, perhaps also the personal and household preferences. General developments in society may also change (leisure) preferences directly. Individualisation, for instance, may decrease preferences for group or traditional family holidays. Finally, general developments and government policies, may change supply conditions. The relaxation of Dutch opening hours in the 1990s, for instance, introduced the possibility to go (fun)shopping on Sundays. Which of these changes will be accommodated by the MERCIN-system will depend on operational and data considerations that are detailed in the chapters 5, 6 and 12.

2.5 Conclusion and discussion

Research on the tourist decision-making process has typically focused on tourist destination choices. Recently, however, it has been recognised that the tourist travel decision also involves choices on other trip dimensions such as the choice of the travel party, the accommodation type, the trip duration, timing, and budget considerations. Also, (facets of) subsequent tourist trips are assumed to influence each other. Moreover, the whole process is said to be dynamic and there is a continuous interaction between the elements of the decision-making process. In this process, decisions made in an earlier phase will condition those made later. Finally, it has been argued that trip choices are the result of complex interactions between preferences, opportunities and constraints regarding the various facets of the tourist decision.

Building on this more comprehensive conceptualisation of tourists' travel decisions, this chapter discussed the conceptual framework for tourist trip patterns underlying the *MERLIN*-system. This framework represents the decision-making context for tourist trip patterns and comprises the interaction between preferences and constraints in the tourist decision-making process, the interrelationship between subsequent trips and the interdependencies between day- and overnight tourist trips. The next chapter discusses modelling approaches that have been developed to quantify tourist choices, followed by a review of existing simulation models in recreation and tourism research in chapter 4. Based on these reviews, chapter 5, finally, will outline the contours of the *MERLIN*-system, including the modelling approaches that are used to model the (facets of) tourist trips and their interactions.

3 Modelling Tourist Choice Behaviour

This chapter reviews the theories and modelling approaches that have been applied to tourist and recreation choice behaviour. Based on this review, the most relevant approaches for the various facets of the tourist decision-making process are selected.

3.1 Introduction

The previous chapter discussed conceptual models of tourist choice behaviour. These descriptive studies provide important insights into tourist choice behaviour and trip patterns. However, many of the studies are less useful for evaluating the effects of policy measures and general developments in society on tourist trip patterns because they are not able to predict tourist choice behaviour and the resulting tourism demand. Prediction requires greater attention to the functional relationship between recreation and tourist behaviour and its determinants (Stynes & Peterson, 1986). Various modelling approaches have been developed since the 1960s to quantify the outcomes of (tourist) choice behaviour. First, *non-behavioural approaches* such as gravity and time-series models were advanced. These models aim to relate aggregate flows of tourists between regions or countries to variables such as the distance between the origin and destination, the associated travelling costs (time and money), the attractiveness and/or price level of the destination, the attractiveness and/or the price level of competing destinations, and characteristics of the origin region such as the number of inhabitants and/or the income level.

In later years, the research community became aware of the fact that the popular gravity and time-series models from the 1960s and 1970s lacked an underpinning consumer theory for their functional form and specification (Lieber & Fesenmaier, 1984; Peterson *et al.*, 1983). This paved the way for the development of different types of *cognitive-behavioural choice models* that are based on (variants) of a conceptual model that explicitly relates choice behaviour to the environment through consideration of perceptions, preference formation and decision making (Timmermans & Golledge, 1990). Typically these behavioural models aim to quantify and test the functional relationship between the characteristics of the decision maker and/or the (characteristics of) tourist choice alternatives (e.g., destinations, transport modes, activities etc.), the preferences for these alternatives, and the probability that an alternative will be selected (Dellaert, 1995). Once a model has been estimated and validated, the consequences of alternative scenarios and policy measures can be assessed by expressing the changes in terms of the independent (or condition) variables of the model, and then using the estimated relationships to predict the most likely behaviour under the assumption of time-invariant behaviour (Timmermans, 2000).

In the field of tourism research, a range of simple to quite sophisticated behavioural models have been applied over the last 15-20 years in response to the increased (appreciation of the) complexity of tourist and travel behaviour. Basically, the existing choice models can be classified into two groups of models. The first group, utility-maximising models, assume individuals to always select the (set of) alternative(s) that maximises their total utility (Ben-Akiva & Lerman, 1985). In, contrast, the second group of models does not make this assumption for a variety of reasons. Advocates of this second group argue, for instance, that it is not realistic to assume that individuals have full-information of all available choice alternatives or that individuals use heuristics and arrive at sub-optimal solutions. Also, they challenge the idea that individuals are free to choose the alternative that matches their preferences best.

The aim of this chapter is to review these behavioural choice modelling approaches. First, the more traditional choice modelling approaches based on utility-maximising behaviour are discussed. Next, the more recently developed non-utility-based models are considered. Finally, these approaches are compared. Based on this review, research question 2 is answered: "How can each facet of the tourist decision-making process be modelled best? And how can the relationships between these facets be included in these empirical models?"

3.2 Utility-based choice modelling approaches

The notion of utility-maximising behaviour presumes individuals to act perfectly rational and consume those products and services (and/or activities) that will maximise the utility derived from their choice behaviour. Two groups of utility-maximising models can be distinguished. First, in the so-called *allocation models* based on *micro-economic consumption theory*, the choice alternatives consist of bundles of commodities, each of which is characterised by a particular price. Based on the premise of "more is better", the rational consumer aims to maximise the utility derived from obtaining (bundles of) commodities within the available financial resources. This is a *continuous* optimisation problem and the solution typically represents the optimal (desired) quantities of each commodity given the prevailing set of (financial, time or other) constraints. Time and money allocation models based on micro-economic theory provide insights into the trade-offs that are made between these constraints. These models do not, however, include information on (the effects of) other personal and household constraints and opportunities, such as age or the presence of leisure commodities, on the trade-offs under considerations[6], whereas these socio-demographic characteristics of the individuals and/or households have proved their worth in describing tourist choice behaviour (Brouwer *et al*., 1994). As a consequence, these models are not very comprehensive because they only model time and money allocation under budget constraints (Ettema, 1996). This may explain the absence of this type of models in recreation and tourism research.

The second group of utility-based models, the so-called disaggregate *choice models*, are also based on micro-economic theory and typically describe how individuals select one alternative from a choice set of alternatives. In these models, mutually exclusive choice alternatives are represented as bundles of attributes and the choice process is *discrete* in nature. These models assume individuals to arrive at some choice by cognitively integrating the utilities they attach to attribute levels representing choice alternatives and then implementing some decision rule. In contrast to time and money allocation models, choice models have increasingly been used to explain tourist and recreation choice behaviour since the 1980s (e.g., Crouch & Louviere, 2000; Dellaert *et al*., 1997; Feather *et al*., 1995; Morley, 1994; Morey *et al*., 1991; Louviere & Timmermans, 1990b; Stynes & Peterson, 1986; Peterson *et al*., 1983). This section therefore only reviews the choice modelling approaches based on the utility-maximising theory models, and in particular their application to tourism and travel behaviour research.

3.2.1 Theory and application of choice modelling approaches

In essence, disaggregate choice models describe cases in which the individual selects one alternative out of a *choice set* of alternatives. These choice models also draw from micro-economic theory. However, since only one alternative is selected from the choice set, Lancaster (1966) suggested that, if utility cannot be derived from the amount of each product consumed, it can be derived from the characteristics, or attributes of the distinct alternatives. Hence, each alternative is characterised by its attribute levels. In formula, the utility of alternative i (U_i ; i = 1,..., I) is written as:

$$U_i = U(X_{1i},...,X_{ki},...,X_{Ki})$$

[6] Eliasson and Mattsson (1998) do specify a stochastic destination attractiveness w_j (in the trip utility function) that varies across individuals, but this parameter is not related to the characteristics of the individual.

where X_{ki} is the *k*-th (k = 1, 2, .. , K) attribute of *i*.

Utility-based choice models assume that choice behaviour is probabilistic in nature. In other words, when faced with a particular choice situation, people will select a choice alternative with a particular probability rather than always selecting the same alternative. The probability of selecting each alternative is related to the utility people expect to derive from an alternative. Drawing upon Luce´s (1959) strict utility theory, *strict utility models* assume the probability of an alternative *i* in choice set *A* ($P(i|A)$) to equal the ratio of the utility associated with that alternative to the sum of the utilities for all alternatives in the choice set. In formula:

$$P(i|A) = \frac{U_i}{\sum_{i'} U_{i'}} \qquad \forall\ i, i' \in A$$

Conversely, the more commonly applied *random utility models* draw upon Thurstone´s (1927a, 1927b) random utility theory, a behavioural theory that recognises that preferences comprise both deterministic and stochastic (random) elements. Choice models based on random utility theory thus assume the measurement of the individual's (unobservable) utility for a choice alternative *i* to consist of an observable structural or deterministic component (V_i) and an error component (ε_i). In formula:

$$U_i = V_i + \varepsilon_i$$

The interpretation underlying the error component is that random variation in the model can be caused by different sources including measurement errors, variations or disturbances in perceptual functions, unobserved influences (attributes) in the environment and unobserved taste variation between individuals (Ben-Akiva & Lerman, 1985). Random utility theory assumes the probability of choosing some choice alternative *i,* is expressed as the probability that the utility associated with that choice alternative exceeds that of all other choice alternatives *i'* in choice set *A*. Formally, this can be expressed as:

$$P(i|A) = P(U_i > U_{i'}) = P(V_i + \varepsilon_i > V_{i'} + \varepsilon_{i'}) = P(\varepsilon_i - \varepsilon_{i'} > V_{i'} - V_i) \qquad \forall\ i' \neq i \in A$$

In order to establish a functional relationship between the choice alternatives and the observed choices, one has to (1) operationalise the structural component (V_i);

and (2) explicate the distributional assumptions regarding the random error component (ε_i). With regard to the structural component, often a linear-in-the-parameters regression-like form is assumed, such as (Crouch & Louviere, 2000):

$$V_i = \beta_{0i} + \sum_k \beta_k X_{ki} + \sum_m \gamma_m Z_{mi}$$

where the β_{0i}, β_k and γ_m comprise the model parameters that capture the base utility of alternative *i* and the effects of the characteristics of the choice alternatives (X_{ki}; as before) and/or the decision maker (Z_{mi}; m = 1, 2, .. , M (the γ_m's are alternative specific parameters)). This specification can be generalised to also include non-linear and non-additive effects such as interactions of *Z*'s with model intercepts or *X*'s (Crouch & Louviere, 2000).

The most common assumption in choice modelling with regard to the random components is that the error terms are independently and identically distributed (IID) according to a Gumbel (Weibul, Type I Extreme Value) distribution[7], resulting in the familiar *multinomial logit (MNL-) model* (Ben-Akiva & Lerman, 1985; McFadden, 1974). Under these assumptions, $\varepsilon_{i'}$ - ε_i follows a logistic distribution and the probability of selecting *i* from the choice set *A* is expressed as:

$$P(i|A) = \frac{\exp(V_i)}{\sum_{i'} \exp(V_{i'})} = \frac{\exp(\beta_{0i} + \sum_k \beta_k X_{ki} + \sum_m \gamma_m Z_{mi})}{\sum_{i'} \exp(\beta_{0i'} + \sum_k \beta_k X_{ki'} + \sum_m \gamma_m Z_{mi'})} \quad \forall\ i, i' \in A$$

Given the assumptions, there are a number of general approaches to estimate model parameters that are efficient, i.e. that are unbiased while no other unbiased estimator has smaller variance. The most commonly used estimation methods include least squares and maximum likelihood estimation techniques.

The popular MNL-model has not been exempted from criticism. In essence, these critiques boil down to (1) the limitations of the (revealed) data that are used to calibrate the MNL-model; (2) the Independence from Irrelevant Alternatives (IIA-) property that follows from the assumed IID distribution of the error terms; and (3)

[7] The use of this distribution can be explained by the fact that choice processes are aimed at selecting the most attractive alternative rather than the selection of the most average alternative in the choice set. This implies that an extreme value distribution like the Gumbel distribution is more appropriate than the average value distribution of, for instance, the Normal distribution (Leonardi & Papageorgiou, 1992). The popularity of the Gumbel distribution should, however, probably mainly be ascribed to its practical properties (Dellaert, 1995, p. 38).

the single choice axiom that restricts the MNL-model to one facet of the decision-making process. These model limitations have often stimulated new lines of research that aim to relax the restrictive assumptions of existing models (Timmermans & Golledge, 1990). The restrictions of the simple MNL-model will be elaborated briefly hereafter.

REVEALED AND STATED CHOICE DATA

Initially, the MNL-model was used as a *revealed preference* (RP) approach explaining observed choice patterns in terms of an underlying utility function. Observations based on real world choices have three major drawbacks. First, using a RP-approach, one cannot obtain data on choice alternatives presently not available in real-world (Louviere & Timmermans, 1990a). Second, the attribute levels of real-world choice alternatives may be highly correlated (intra-attribute correlations), thus providing false information regarding the real trade-offs due to the resulting problems of multicollinearity and lack of variability (Morley, 1994; Ben-Akiva & Morikawa, 1990). Third and finally, RP-data may not always include information regarding the choice set considered by the decision maker.

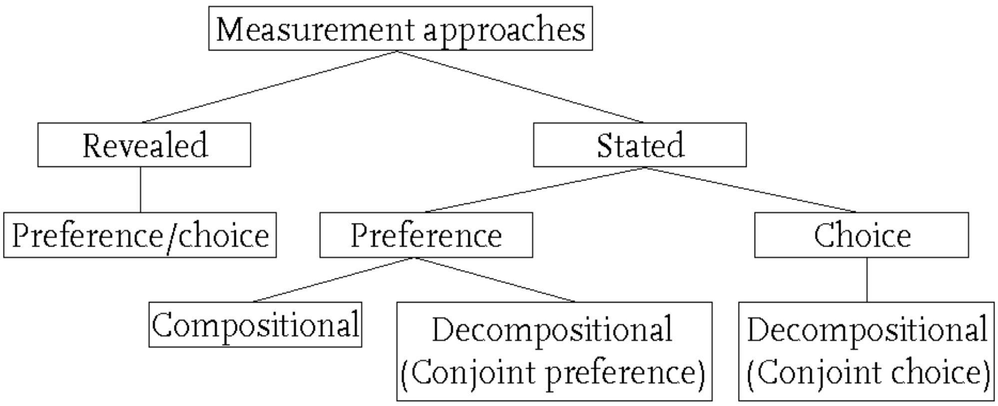


Figure 3.1 An overview of measurement approaches (Kemperman, 2000; p. 83)

In response to these limitations, new measuring approaches have been developed. These so-called *stated measurement* approaches have the advantage of being based on experimental measurements: respondents are presented with hypothetical choice alternatives in carefully designed settings which allow the researcher to control for the above mentioned problems (Louviere & Timmermans, 1990a, 1990b). Within these experiments, both preferences and choices can be observed. *Stated preferences for attributes* of hypothetical choice alternatives can be observed by requesting respondents to evaluate attribute levels and the importance of each attribute on some scale. Using these data, the total utility of an alternative can be *composed* by multiplying the attractiveness of attribute levels and the importance scores (under the assumption of some choice rule). *Stated preferences for alternatives,* on the other hand, can be observed by requesting respondents to rate or rank-order hypothetical choice alternatives. *Stated choices,* finally, are observed when the respondent makes a discrete choice between hypothetical choice alternatives. Using the latter preference or choice data, the total utility for these choice alternatives can be *decomposed* into the part-worth utilities for each attribute level; this is sometimes also referred to as *conjoint analysis*. In comparison to revealed and conjoint preference analysis, stated choice experiments have the advantage of allowing the estimation of both a preference or utility function *and* a choice model. Figure 3.1 provides an overview of these measuring approaches.

THE INDEPENDENCE FROM IRRELEVANT ALTERNATIVES (IIA)-PROPERTY

As the MNL-model assumes the error terms to be IID, it implicitly assumes that the variance of all error terms are equal, and that there exist no systematic correlation between them. As a consequence, the MNL-model also assumes that the relative odds of choosing a potential alternative depend only on their measured attributes and is not affected by the composition of the choice set. Due to this so-called *Independence from Irrelevant Alternatives* (IIA) property, the MNL-model is unable to account for substitution and dominance effects. An example in tourism would be to introduce a zoo into a region that includes a museum, an amusement park and a zoo. Given this change, the MNL-model would not be able to account for the similarities between the existing and the new zoo, and predict the new zoo to equally draw visitors from the three existing tourist attractions.

In response to the IIA-limitation, many alternative models have been developed such as nested and random-parameter logit, probit and negative exponential distribution, and (generalised) extreme value models. In addition, background and cross-effects have been included (Bunch, 1991; Timmermans *et al*., 1991; Timmermans & Golledge, 1990; Daganzo, 1979; McFadden, 1978). Application of these alternative models in recreation and tourism can be found in Train (1998), Dellaert (1995), and Borgers *et al*. (1987).

THE SINGLE CHOICE AXIOM

Traditional discrete choice models only consider single choices such as the choice of transport mode *or* holiday destination. However, in chapters 1 and 2, it was argued that the choices involved in bringing together the elements of a tourist trip typically are multi-facetted choice processes. Non-IIA models such as nested logit and probit models, have often been used to model simultaneous and sequential decision processes of multiple choices as they allow for choices to be hierarchically structured. An application to urban tourism research, for instance, would be tourists' choice of destination *and* transport modes for short city breaks (see Figure 3.2).

Dellaert *et al.* (1997) use the term *portfolio choice* to indicate choices between combinations of alternatives and propose a general approach to conjoint choice models for portfolio choices. Application to urban tourism indicate that models of tourists' purchases of separate services may severely overestimate the influence of policy actions on demand if they are in fact purchased in a portfolio combination with other services (Dellaert *et al.*, 1997; Dellaert, 1995).

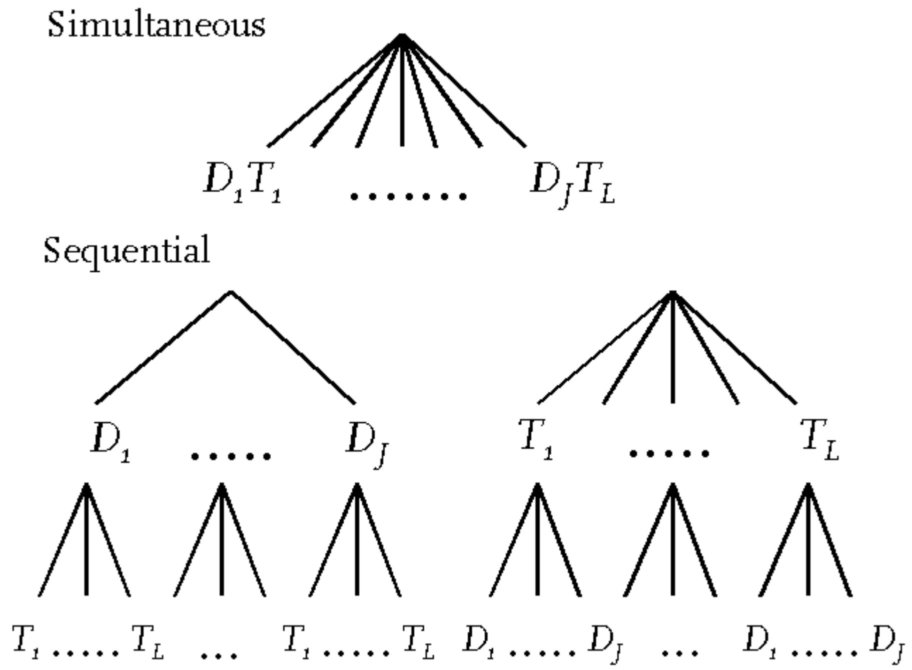


Figure 3.2 Choice structures for combinations of destination choice alternatives ($D_1,..,D_J$) and transport mode alternatives ($T_1,..,T_L$); based on Dellaert (1995)

Since a hierarchical model is used to model part of the behaviour under consideration in the MERCIN-system (this decision is accounted for in section 3.4), the nested-logit model will be discussed in more detail using the combined destination-transport mode choice as an example (see the "sequential" choice structures in Figure 3.2; this example discusses the structure on the left-hand side of the figure, where transport mode choices are made conditional upon the destination choice; the discussion and references are based on Dellaert (1995)).

Given a portfolio choice of two elements, the error component of the combined destination-transport mode choice is divided into two independent error terms. In formula:

$$U_{D_jT_l} = V_{D_j} + V_{T_l} + V_{D_jT_l} + \varepsilon_{D_j} + \varepsilon_{D_jT_l}$$

where U_{DjTl} represents the total utility of the combined destination-transport mode choice, and V_{Dj}, V_{Tl} and V_{DjTl} are the structural utilities of the (joint) alternatives. The first error term (ε_{Dj}) represents the disturbances of the destination choice in the highest nest, while the second (ε_{DjTl}) represents the disturbances in the joint destination-transport mode choice. As a consequence, the differences in the unobserved error components of choice alternatives that share elements - in this example the alternatives within a particular destination choice nest - are smaller than the differences in the error terms of alternatives that do not belong to the same nest of destination choice alternatives.

The disturbances in the joint destination-transport mode choice are assumed to be IID Gumbel distributed with parameters (0, μ^{low}), where μ^{low} is the scaling factor of the lower level. The disturbances in the destination choice at the highest level (ε_{Dj}) are assumed to be IID Gumbel distributed, but with parameters ((1/μ^{low}) ln $\Sigma_{l\in L}$ exp((V_{Tl} + V_{DjTl}) μ^{low}), μ^{low}) (Johnson & Kotz, 1970). Given these assumptions, the probability that a destination D_j will be selected is expressed as:

$$P(D_j) = \frac{\exp((V_{D_j} + V_{\max\{D_jT_l\}})\mu^{high})}{\sum_{j'} \exp((V_{D_{j'}} + V_{\max\{D_{j'}T_l\}})\mu^{high})} \qquad \forall \; j, j' \in \; J$$

where μ^{high} is the scaling factor of the higher level and $V_{\max\{DjTl\}}$ represents the maximum utility of the lower level attributes of the alternatives in the set {D_jT_1, ... , D_jT_L }. This is also called the Inclusive Value (IV) of the nest. The probability that transport mode T_l is selected conditional upon destination choice alternative D_j is expressed as:

$$P(T_l | D_j) = \frac{\exp((V_{T_l} + V_{D_jT_l})\mu^{low})}{\sum_{l'} \exp((V_{T_{l'}} + V_{D_jT_{l'}})\mu^{low})} \qquad \forall \; l, l' \in \; L$$

It is not possible to estimate the two scale parameters (μ^{low} and μ^{high}) from the data. In practice, therefore the ‘low’ scaling factor of joint destination-transport mode choice is arbitrarily set to 1 (μ^{low} = 1) so that (the estimation) of the scale of the higher level destination choice represents the ratio of the two scales (μ^{high}/μ^{low}). If the hierarchical structure adequately describes the data, this parameter for the scale differences takes on a value between zero and one. (In addition, if there are three or more hierarchical levels, the values of the scale parameters should ascend in the direction of the lower level (Ben-Akiva & Lerman, 1985)). If the scale parameter equals one, the model reduces to the joint logit model, implying that no systematic correlations exist between any pair of alternatives that are identical at a particular dimension (Dellaert, 1995; Ben-Akiva & Lerman, 1985). If it equals zero, on the other hand, the model reduces to the simple MNL-model at the higher level and the choice at the lower level does not affect the higher level choice at all.

3.2.2 Conclusion and discussion

This section reviewed models that are based on the assumption of utility-maximising behaviour. First, *time and money allocation models* based on micro-economic theory were mentioned briefly. These models aim at explaining the optimal allocation of time and money budgets and provide insights into the trade-offs that are made between these constraints. Despite some apparent advantages of these models, they have not been applied in the field of recreation and tourism.

In contrast, the second group of utility-maximising models, the so-called *disaggregate choice models*, have been applied to tourist choices rather frequently. One of the most popular choice models is the *Multinomial Logit* (*MNL-) model*. Based on random utility theory, this modelling approach relates observed choice behaviour to the characteristics of the choice alternatives and/or the decision maker by assuming the utility derived from the choice alternatives to consist of a deterministic and a stochastic or random component. Given several assumptions regarding the structure of the deterministic utility component and the distribution of the error term, the part-worth utilities of the levels of the explanatory variables are estimated.

Despite the popularity of utility-maximising choice models, it has been argued that the theory of utility maximisation may not represent individual decision-making very accurately. In particular, it is unrealistic to assume that decision makers possess adequate knowledge (and willingness) to evaluate the utility of each alternative in the choice set, identify the alternative that gives the highest utility, and then select that alternative. In other words, people may not always behave rational. Also, it may be unrealistic to assume that people are unhindered in their choices. And although time and money allocation models assume individuals to be constrained by financial and temporal budgets, there may be many more sources of constraints, including coupling constraints emanating from the individuals' social situation or authority constraints related to institutional regulations. Finally, given the apparent desire for variation in leisure activities, it is not realistic to assume individuals repeatedly select the same choice alternative when faced with the same set of opportunities (Kemperman *et al.*, 2000; Van der Heijden & Timmermans, 1987; Fesenmaier, 1985). In response to these comments, therefore, various alternative theories and modelling approaches have been developed that do not assume individuals to always act rationally. The next section discusses these approaches.

3.3 Non-utility-maximising modelling approaches

Advocates of non-utility based approaches challenge the assumption of utility-maximising theory that decision-makers are able to arrive at optimal solutions. The *constraint theories* that were discussed in the previous chapter, for instance, argue that people are not free to choose the alternatives they prefer because they are restricted by limited time and money budgets or situational constraints, such as the weather. In response to the preference-based utility-maximising theory, the traditional view was that constraints, when present, would intervene on preferences and restrain people from participating in the activities they desired. Later, more sophisticated theories and models regarding the relationship between preferences and constraints were advanced. In this more comprehensive view, constraints may act differently on tourist preferences depending on their source, their intensity, and the stage of the decision-making process during which they are negotiated (Jackson *et al.*, 1993). In addition, different coping (or negotiation) strategies have been reported (Kay & Jackson, 1991).

The interest of constraints research in coping strategies and the negotiation of preferences and constraints in the decision-making process reminds one of approaches that aim to imitate the working of human brains or human decision-making processes. However, based on modern psychological and physiological theory, the latter approaches challenge the assumption of utility-maximisation for different reasons. In particular, they dispute the assumption that people are able to arrive at optimal solutions because decision-makers have imperfect information and/or use sub-optimal choice processes.

In contrast to the quantitative utility-based approaches, constraint-based models and models based on psychological and physiological theory are more qualitative approaches. As a consequence, the latter approaches offer more flexibility in modelling tourist choice behaviour because they do not assume an a-priori functional form nor do they require variables to follow a particular distribution (such as the IID Gumbel assumption for error terms in the MNL-model). In addition, both compensatory and non-compensatory decision rules can be included[8]. The development of models that aim to imitate the working of human brains or human decision-making processes is often conceptualised as a problem of training a (rule-based) system based on examples, i.e. observed trip patterns. These so-called inductive (or empirical) learning processes have received ample attention in the machine learning literature (Arentze & Timmermans, 2000a). Within these cognitive approaches, two paradigms can be distinguished, the *connectionist* and the *symbolic search space paradigm*. The remainder of this section reviews the modelling approaches that were developed from these paradigms (constraint-based approaches will not be reviewed separately because constraints are often included within other non-utility based approaches either explicitly or otherwise).

3.3.1 Theory and application of Neural Networks

The *connectionist paradigm* advocates models based on a direct analogy to the human brain in terms of how people store and recall information. Within this approach, operational models are known as artificial *neural networks* (NNs). NNs are conceptualised as a set of processing nodes that are interconnected by weight factors. The architecture of a NN (see Figure 3.3) typically includes an *input layer* (in terms of tourist choices: a data matrix which represents the conditions of the choice situations), one or more *hidden layers*, and an *output layer* (in terms of tourist choices: revealed or stated trip choice behaviour). Each layer can have many processing nodes (or neurons), that are interconnected by weights (W_{ij}). A NN is trained by processing cases. In this process, the input layer receives stimuli which causes the weights between the nodes to be adjusted (i.e., the network "learns"). In uni-directional *feed-forward* networks, nodes can learn only from preceding layers. In *feedback* networks, on the other hand, recall is bi-directional where nodes are also allowed to feedback to later layers. This learning is *supervised* if the process is defined by pre-specified output categories (as is the case with discrete choice models); it is *unsupervised* if the network has to organise the input-data by computing similarities and differences between data-points without a pre-specified desired output.

[8] Although in principle (approximations to) non-compensatory decision processes can be formulated in quantitative models, virtually all utility-based models of tourist choice behaviour assume that tourists will balance positive evaluations on one aspect of the decision situation against negative evaluations of other aspects.

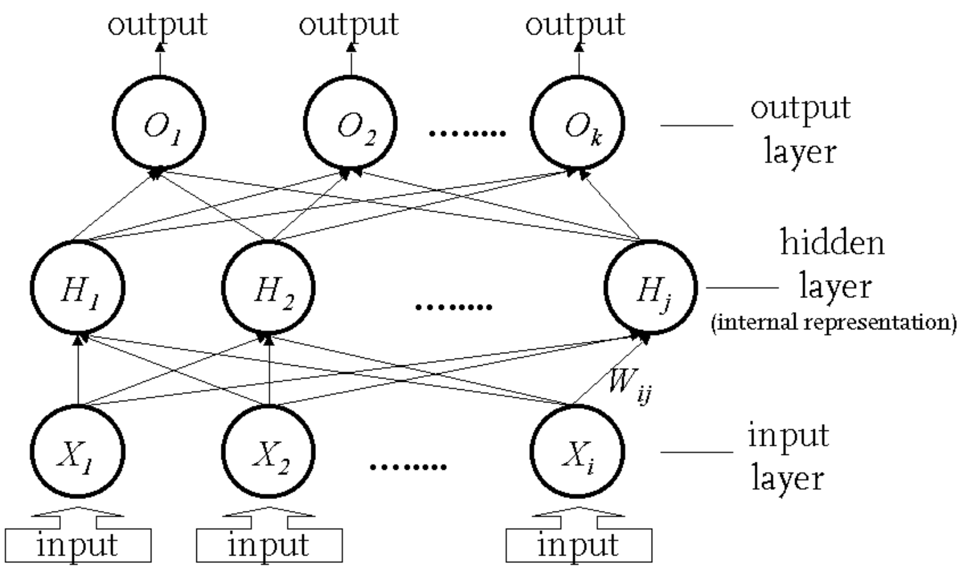


Figure 3.3 The general architecture of a neural network

Once a NN has been trained it can be used to perform forecasting (Pattie & Snyder, 1996), classification (Mazanec, 1999, 1992), data reduction and optimising tasks. In the field of tourism (and other marketing sciences), however, the majority of papers are still focussed on comparing NNs with established methodologies (Ganglmair & Wooliscrof, 2001; Davies *et al.*, 1999; Law & Au, 1999; Uysal & El Roubi, 1999; Jeng, 1995). According to Mazanec and Moutinho (1999), this will go on until more reliable conclusions regarding the use and applicability of this approach emerge.

3.3.2 Theory and application of computational process models

The second approach to developing models that imitate human decision-making processes is the *symbolic search space paradigm*. In contrast to connectionist theory, this paradigm adopts an analytical approach to problem solving by assuming a highly structured representation of the choice situation, the search space, that is often represented by a tree structure (Ettema, 1996). The search process is assumed to consist of efficiently moving through this space of states to reach the most appropriate solution. From this tradition, *production systems* or *computational process models* have been developed. These *rule-based approaches* assume individuals to use heuristic decision-making rules that represent decisions that have worked out satisfactorily under similar conditions in the past. Decision rules reflect both desires and constraints of decision-makers and are represented by logical expressions (e.g., IF <condition state(s)> THEN <action>). Rules can both be *crisp*, where all cases complying with the specified conditions are assigned to the most likely choice alternative (*deterministic all-or-nothing assignment*), or rules can be more *fuzzy* and/or *probabilistic*, where cases complying with the specified conditions (with a certain level of imprecision or fuzziness) are assigned to the available choice alternatives with a certain probability (*probabilistic assignment*).

The use of rule-based models brings about two important considerations. First, heuristics have to be represented using a particular formalism. The challenge in developing a rule-based system is to select a formalism that guarantees the consistency, exclusiveness and exhaustiveness of the rulebases, that lends itself for theoretical interpretation, and that supports the use of learning mechanisms to derive rules from observed tourist trip patterns (Arentze & Timmermans, 2000b; Wets, 1998; Witlox, 1998). Production systems (i.e. sets of IF-THEN-rules (or productions)), decisions plan nets, decision tables and decision trees have been used for this purpose. Due to the strict format, decision tables[9] (DTs) offer the most compact, efficient and effective visual presentation, ease of manipulation and ability to check information on consistency, exclusiveness and exhaustiveness (Wets, 1998; Arentze *et al*., 1996; Witlox, 1995; Vanthienen, 1994; Palvia & Gordon, 1992). *Consistency* means that for each possible combination of condition states, it should be unmistakable which actions should be performed. In addition, *exclusivity* implies that at least one element of the condition part in a decision rule does not intersect with the corresponding element in the condition part of another rule. *Exhaustivity*, finally, means that every condition state of each condition variable is included in the rulebase, and that for each combination of condition values at least one action is specified.

Table 3.1 Example of a decision table

Av. temperature during high season	$T < 20^0$ C	$T \geq 20^0$ C	$T \geq 20^0$ C
Rain season	-	no	yes
Visit?			
No	X		X
Yes		X	
Rule number	1	2	3

To demonstrate these logical requirements, table 3.1 presents an example of a deterministic decision table comprising two condition variables and two actions:

[9] A decision table (DT) is a matrix-like representation of the decision making process that consists of *condition* variables (left upper part), their levels or *states* (right upper part), *actions* or decisions (left bottom part) and *rules* that link condition states to actions (right bottom part) (Wets, 1998; Witlox, 1995; Verhelst, 1993). DTs can also be represented in a *tree*-like formalism.

"visit a particular destination" (Yes) and "do not visit a particular destination" (No). This DT is consistent because for each combinations of the four condition states the actions are clearly indicated: when the average temperature is lower than 20^0 C, the destination is not visited, whether there is a rain season or not (rule #1); when the average temperature exceeds 20^0 C, and there is no rain season, the destination is visited (rule #2); finally, when the average temperature exceeds 20^0 C, and there *is* a rain season, the destination is not visited (rule #3). The DT is exclusive because no combination of temperature and rain season can categorised into two or more rules at the same time. The DT is exhaustive, finally, because all possible combinations of the four condition states are captured in either one of the three decision rules.

The second consideration when using rule-based models is that the decision rules have to be derived. Initially, for lack of statistical techniques to derive rules from data, rulebases were obtained from qualitative techniques such as expert interviews and think-aloud protocols (Arentze *et al*., 2001). However, these techniques lack the ability to test the validity and the predictive power of the derived rulebase. Only recently, algorithms originating from information theory and statistics have been applied to induce decision rules from empirical data. Basically, two approaches can be distinguished. The first approach is to induce decision rules *directly from empirical data*. Examples include approaches based on the rough set theory, and Genetic Algorithms (Oliver, 1994; Goldberg, 1989; Greene & Smith, 1987). The disadvantage of these algorithms, however, is that they run the risk of inducing incomplete sets of decision rules. In other words, the induced sets of decision rules may not be exhaustive and they will not necessarily be able to classify all possible cases that are not part of the data set from which the rules were derived. Also, these sets may contain conflicting rules, i.e. the decision rules are not always exclusive (the latter problem can be solved by giving priority to the rule with the highest fitness).

Alternatively, several algorithms build decision *tree-structures* from empirical data, and subsequently transform this tree into a set of rules. By considering splits and mergers as the only permissible operations, these algorithms make sure that the sets of tourist decision rules are exclusive, exhaustive and consistent. In the context of tourist decision-making, these tree-induction systems use condition variables to repeatedly partition the sample of observed tourist choices into mutually exclusive and exhaustive sets of conditions states that are more homogeneous with regard to the tourists' decisions (or actions). Many different criteria can be defined for selecting the best split (i.e., best (combination of) condition states), and this recursive process repeats itself until some pre-specified stopping criterion is satisfied. Stopping criteria often include significance or improvement testing of possible combinations of condition states and/or the specification of a minimum number of observations within each (set of) condition states before and/or after split. The sets of observed tourist choices are thus defined by combinations of (i.e., interactions between) condition states (Strambi & Van de Bilt, 1998; Magidson, 1995). By linking the response distribution of a set of observed tourist choices defined by a particular set of condition states to the actions, a decision rule is obtained. The most commonly applied tree-building algorithms are C4.5 (Quinlan, 1993), CART (Breiman *et al*., 1984) and CHAID (Kass, 1980).

Although it has been suggested by several authors that travellers use heuristics, or choice rules, to set priorities for their choice decisions (King & Woodside, 2001; Bervaes *et al*., 1996; Bronner & De Hoog, 1985), operational rule-based models of tourist and recreation behaviour are rare. An exception includes two studies published in 2000 by Law and Au, that, based on the rough set theory, induce two groups of decision rules to predict tourist shopping respectively sightseeing expenditures as percentage of the total expenditures in terms of (very) high and low (Law & Au, 2000; Au & Law, 2000). Belonging to the group of algorithms that induce rules directly from a given data set, however, this approach does not produce exhaustive sets of decision rules. In both of the testing data sets, for instance, one of the 17 cases could not be classified using the induced set of rules.

3.3.3 Conclusion and discussion

This section reviewed non-utility based approaches to tourist choice behaviour. These approaches challenge the assumption that people are able to arrive at optimal solutions because they are restricted by constraints, including imperfect information, and/or use sub-optimal choice processes. The development of qualitative non-utility-based models is often conceptualised as a problem of training a system based on examples. Two approaches were discussed. First, *neural networks* (NNs) aim to imitate the working of human brains. The advantages of NNs are that they are able to cope with the continuous influx of data, nonlinearity, interactions and missing data; moreover, they are able to address high-dimensional problems that are computationally intractable for conventional methodology and they have been reported to produce more stable results. However, these advantages come at the cost of the lack of significance testing and the ‘black-box’ nature of the relationship between input and output data. The latter disadvantage is due to the use of subsymbolic units (i.e., the network nodes) that can be either activated of non-activated. This conceptualisation cannot always be related unambiguously to concepts in the real choice situation. Hence, the contribution of NNs to our understanding of causal relationships between choice conditions (i.e. preferences, opportunities and constraints) and the outcome of the decision-making processes may be limited[10].

Rule-based approaches can also cope with non-linearity and missing data. Compared to NNs, however, rule-based approaches have the advantage of producing relations between input and output data that are relatively easy to interpret. Also, some rule-inducing algorithms include significance testing tools at the level of merging or splitting sets of conditions states. Furthermore, rule-based formalisms, decision tables in particular, offer a useful visual presentation that allows for tests on the completeness, correctness and consistency of the model. However, the algorithms that produce sets of tourist decision rules that are exclusive, exhaustive and consistent (i.e., the tree-building algorithms), are often only one-step optimal and not overall-optimal because each split is optimised separately and independent from the possible splits that might follow. In addition, they are often restricted to one dimension of the travel decision only. As a consequence, multi-dimensional choice situations like the travel decision can only be modelled using a series of sequential rule-based models in which, at best, outcomes of previous decisions are included as conditions.

3.4 Utility- vs. non-utility-based behavioural choice approaches

The previous two sections discussed utility-based and non-utility-based approaches to tourist decision-making. Some of these approaches, utility-based (discrete) choice models in particular, are common to recreation and tourism research. In contrast, non-utility-based models have only been introduced to the field rather recently. The purpose of this section, then, is to assess these approaches in the light of the research question 2. In other words, *how can each facet of the tourist decision-making process be modelled best?* And, *how can the relationships between these facets be included in these empirical models*? Based on the conceptual framework presented in the previous chapter, two substantially different phases of the tourist decision-making process can be distinguished. First, there is the problem of

[10] Recently, hybrid approaches have been advanced to relieve this drawback. By extracting symbolic rules from trained NNs, so-called Knowledge-Based Neural Networks (KBNNs) aim to exploit the complementary properties of knowledge-based and neural network paradigms to obtain more powerful and robust systems (Taha & Ghosh, 1999).

allocating a certain amount of (time and money) resources to tourist trips. Given the set of tourist trips an individual wants to pursue during a year, the next step is to characterise each tourist trip by its profile, including the interrelated choices of facets such as the travel party, the destination, and so on. The remainder of this section considers the appropriate modelling approaches for each of these two phases in turn.

3.4.1 The selection of the annual set of tourist trips

Based on the conceptual model, the selection of the annual set of tourist trips is conceptualised as an allocation problem in which the decision to pursue tourist trips is traded-off against other time and money consuming activities. In addition, the time allocated to tourist trips has to be distributed among the distinct categories of trips that differ in intensity and/or resource attachment, i.e. day-trips and overnight holidays. This reminds one of the allocation models based on micro-economic theory that consider this to be a *continuous* process. In contrast to the assumptions underlying these models, however, several studies indicate that it is not realistic to assume that tourist will always allocate all additional financial and time resources to tourist activities (e.g., Dirven *et al*., 1998; Wijker, 1998). In addition, the premise of “more is better” (Kraan, 1996) common to micro-economic allocation models does not apply to tourist trip choices. Also, it has been argued that these models are not able to capture the effects of the socio-demographics that are important to describing tourist choice behaviour.

Alternatively, several studies have emphasised the mixed discrete choice-continuous allocation nature of resource-allocation problems where the optimal discrete choice partially depends on the outcome of the continuous choice. Applying this view to tourist (day- and overnight) trips, tourist time allocation would be a two-stage process in which, first, the (potential) tourist decides whether or not to participate in day- and/or overnight tourist trips - a discrete *participation choice.* And second, conditional on this participation decision, the individual would allocate a certain amount of resources to various tourist trips - a more continuous *trip quantity or frequency choice* process. Several utility-based modelling approaches have been used to models discrete/continuous choice processes, including Tobit models (Yamamoto & Kitamura, 1999; Meloni *et al*., 1998; Kitamura *et al*., 1996) and conditional indirect utility functions (including various assumptions regarding the joint distribution of the random elements ε (Hanemann, 1984)). From a non-utility based approach, a classical CHAID analysis has been used to model trip-generation (Strambi & Van de Bilt, 1998). Unfortunately, however, these models are restricted to one (Strambi & Van de Bilt, 1998) or two activities (e.g., in-home vs. out-of-home activity engagement (Kitamura *et al*., 1996)), assume that resources will be allocated to all activities (this reduces the model to a continuous resource allocation problem; Bhat & Misra, 1999), and/or assume that the consumer selects only one of the discrete alternatives (Hanemann, 1984). In the light of tourist trip choices, none of these assumptions will hold.

In recreation demand modelling, combined continuous/discrete models are used to model the allocation of recreational trips across substitute sites conditional upon the trip frequencies per time period. In most studies, the number of trips is modelled using count-models (i.e., Poisson or negative binomial regression models; Siderelis & Moore, 1998), whereas the site-selection phase is modelled using discrete choice models such as the (nested) MNL-model (Feather *et al*., 1995; Stynes & Peterson, 1986). The disadvantage of count-data models, however, is that they often only describe the number of trips of one type, thus excluding interrelationships between the trip frequencies of the different trip types within the tourist's choice set (e.g., day-trips and holidays). Discrete choice models like the MNL-model, on the other hand, are pre-eminently able to model the preferences for different alternatives in a choice set. In addition, they have the advantage of being able to model discrete choice as well as allocation or rating data. Given the (assumed) two-staged nature of the tourist resource allocation process, it is therefore proposed to model resource allocation to tourist activities using a hierarchical extension of the MNL-model to combined choices[11]. Furthermore, to reduce the complexity of the model, the allocation process is restricted to time resources only. In the hierarchical model, the allocation of time to various tourist trips (the *trip quantity or frequency choice* process) is modelled conditionally upon the decision whether or not to participate in day- and/or overnight tourist trips. Chapter 7 elaborates the structure of this model.

3.4.2 The interrelated choices of facets

Given the annual set of tourist trips, each tourist trip has to be characterised by a profile of choice facets, including the travel party, the timing of the trip, the destination, the accommodation, the transport mode, the date of departure and the expenditures. Except for the latter facet, these choices typically comprise discrete choices from a set of choice alternatives. Traditionally, these choices were modelled

[11] It is acknowledged that this process can also be modelled using neural networks. However, since we are also interested in the relationship between the choice conditions and the allocation of time to tourist trips, this option was rejected.

using utility-based discrete choice models, MNL-models in particular, that elucidate relations between observed choice behaviour on the one hand, and the characteristics of the choice alternatives and the decision maker on the other. It was argued, however, that the more recently advanced non-utility based models offer more flexibility in modelling tourist choice because (1) these qualitative approaches do not impose rigorous assumptions on the functional form of the model and the distribution of the variables; and (2) both compensatory and non-compensatory decision rules can be included. This potential higher flexibility in representing alternative decision processes of qualitative models does not necessarily imply that they are better predictors of observed choice behaviour than conventional quantitative models. However, several studies in shopping location choices (Thill & Wheeler, 2000), daily activity scheduling and transport (Wets *et al*., 2000; Arentze *et al*., 2000) and tourist timing choices (Van Middelkoop *et al*., 2000) indicate that decision tree induction systems at least match the results of conventional utility-based models.

Based on the above considerations, it is proposed to model the choice facets of tourist trips using *non-utility based approaches*. Also, within this group of qualitative approaches, preference is given to *computational process models* because of the clear relationship between conditions and choice outcomes, and the direct availability of statistical tests. Furthermore *probabilistic rules* will be used and *Monte Carlo simulation* will be applied to identify the choice alternative that will be selected. Probabilistic rules are preferred over deterministic rules to preserve the heterogeneity within the rules that are inevitable in empirical situations. Finally, the rulebases will be laid down in *decision tables* because this formalism offers the best visual presentation and allows for test of its consistency, exclusiveness and exhaustiveness. In short, the interrelated choices of facets for each tourist trip will be modelled using a series of sequential probabilistic decision tables that will be simulated using Monte Carlo simulation. The outcomes of previous decisions are included as conditions in subsequent decisions to capture the interdependencies between successive choice facets.

The final decision that remains involves the choice of the appropriate rule-induction algorithm. First, a tree-induction algorithm is preferred because these algorithms produce complete, exhaustive and consistent sets of tourist decision rules. As mentioned before, the three most commonly applied tree-building algorithms are C4.5, CART and CHAID. C4.5- and CHAID-based algorithms have the advantage of allowing multiple-way instead of only binary splits between the condition states of a condition variable. C4.5- and CART-based algorithms have the advantage of producing more concise sets of decision rules because of their built-in pruning options. In addition, in case of continuous condition variables, a C4.5-based algorithm selects and tests the optimal split of a condition variable into discrete categories. A CHAID-based algorithm, on the other hand, has the advantages of built-in significance testing for each split between or merge of condition states in relation to the observed tourist choices. Also, due to the criterion based on the (significance of the) χ^2-statistic, this algorithm takes into account the whole response distribution rather than just the mean and the standard deviation. This is also true for the C4.5- and CART-based algorithms. However the pruning criteria that are used as part of these algorithms, often favour the dominant responses within the set of observations. Given these advantages and disadvantages of the tree-induction algorithms, a CHAID-based algorithm will be used to derive the choice rules for the facets of tourist trips. This decision is motivated by the greater sensitivity of this algorithm to the whole response distribution, which is favourable in the light of probabilistic decision rules. It is acknowledged, however, that this comes at the cost of more elaborate sets of decision rules and that, given the present state of knowledge regarding the application of rule-induction systems in the social sciences, the choice in favour of any algorithm is at least in part arbitrary.

3.5 Conclusion and discussion

This chapter discussed theories and modelling approaches that have been applied to tourist and recreation choice behaviour. First, traditional approaches based on the assumption of utility-maximising behaviour were discussed. One of the best known models originating from these approaches, i.e. the Mulitnomial Logit (MNL-)model, has been applied to recreation and tourism rather frequently. In addition, several more complex utility-based models that relieve one or more of the limiting or unrealistic assumptions of the MNL-model have found their way into the tourism research community. At a more fundamental level, however, advocates of a second group of models challenge the assumption of utility-maximising behaviour for a variety of reasons. Within the non-utility based modelling approaches both the connectionist and the symbolic search paradigm were reviewed in terms of the underlying theories, the modelling approaches, and the application within the field for recreation and tourism. Finally, based on these reviews of modelling approaches that have been applied to tourist choice behaviour, the final part of this chapter identified the modelling approaches that will be used to represent the various phases and facets of the tourist choice process in the *MERLIN*-system. The next chapter will provide the final conceptual building blocks for the development of the MERLIN-system by reviewing the principals underlying *simulation approaches* that allow us to build representations of complex real-world situations by synthesising from these relatively simple parts.

4 Simulation: Principles and Application in Recreation and Tourism Research

This chapter introduces simulation as an approach to analysing complex problems. First, the principles underlying simulation approaches are reviewed. Next, existing simulation models in recreation and tourism research are discussed.

4.1 Introduction

Simulation is the process of designing a model of a real world system and conducting experiments with this model for the purpose either of understanding the behaviour of the system or of evaluating various strategies (within the limits imposed by a criterion or set of criteria) of the operation of the system (Shannon, 1975, p. 2). Stated differently, simulation is a *goal-oriented experiment* that supports the comprehension of a particular problem and that may support the *insight* into and/or *decision-making* with regard to that problem. In essence, therefore, every model or representation of a real world situation is a form of simulation, whether it would be an application of one of the models that was discussed in the previous chapter, computer games like SimCity and Command and Conquer, or a child's attempt to understand its interactions with other humans or objects by playing with toys (Trick, 1996; Fishwick, 1995; Shannon, 1975). In practice, however, the term "simulation" is usually associated with the class of methods to be used when some system of interest is too complicated to be handled "analytically" (Cesario, 1975). This chapter follows this definition.

Restricted to *computer* simulation models, i.e. models adapted for simulation on a computer, and their *application in research and planning*, the structure of simulation models is mainly determined by the relations between dependent (or action) and explanatory (or condition) variables expressed in mathematical or logical terms. These relations thus include both algebraic equations and decision structures that can be characterised as complex "if-then"-relations (Merz, 1991). In contrast to the previously discussed structural models, however, simulation allows analysts to construct approximations of complex real-world situations by synthesising from relatively simple parts (Timmermans, 2000; Shannon, 1975). In this context, the structural models are often used to describe *parts* of the behaviour under consideration in simulation models. The distinction between "normal" and simulation models, therefore, is a gradual one, where generally models increase in complexity and lead time to develop a forecast when one moves from the structural models that were discussed previously to the *systems* or *simulation models* that are the subject of this chapter (Stynes, 1983). Like "normal" models, however, simulation models can be used for both scenario-evaluations and forecasting. *Scenario-evaluations* should give insight in what would happen if particular exogenous variables (e.g. exchange rates, demographics) are manipulated. *Forecasting* should give insight into the future development in the tourist sector given expected developments in the exogenous variables (Van Dijk *et al*., 1991).

The objective of this chapter is to discuss the principles underlying simulation approaches (section 4.2). Also, this chapter reviews applications of the simulation approach in the field of recreation and tourism research (section 4.3). Together these reviews provide the necessary background for the positioning and development of the MERCIN-system.

4.2 Fundamentals of simulation

4.2.1 The art and science of simulation

Computer simulation techniques present the practical opportunity to model reality as closely as possible (Levine & Lodwick, 1983a). From a conceptual point of view, systems or simulation models assume that recreation and tourism patterns result from complex relationships involving both historical factors and present conditions (Stynes, 1983). Practically, this is often accomplished by (1) analysing the entire system (in our case, the tourist decision-making process) and identifying the fundamental components and their relationships; (2) modelling these components and their interrelationships (in our case, the facets of the tourist decision-making process in relation to each other and the characteristics of the individual); and (3) constructing a complex model by synthesising from these relatively simple parts while taking into account the identified interrelationship (Shannon, 1975). Fishwick (1995) refers to the resulting model as a "*multimodel*" that contains multiple integrated models each of which represents a level of granularity of the system under consideration. According to Shannon (1975), simulation models are also called *input-output models*. This means that given a certain user-defined input and the structure of the system (under the conditions specified by the experimenter), the system yields a certain output. In this sense, simulation models are "run" rather than "solved". An simulation *experiment* or *analysis*, then, is a set of runs in which the decision conditions are modified according to a predetermined plan from one set of replicated runs to another (Shechter & Lucas, 1978). When assessing the direct and side effects of a given set of conditions, a *baseline simulation* is needed to describe the 'no change' simulation. Usually, the baseline simulation describes the ‘status quo’ development, though it is a matter of interest and definition as to what will, in fact be the baseline (Merz, 1991). Finally, it is noted that simulation models are *descriptive* rather than prescriptive: they can only tell how the system works under certain conditions, not how to arrange the conditions to make the system work best (Trick, 1996).

Simulation can be a very powerful tool for solving problems because, in contrast to the structural models that where discussed in the previous chapter, it is limited only to the amount of time you want to spend getting data and programming it (Trick, 1996). Simulation should be considered when (Trick, 1996; Fishwick, 1995; Merz, 1991; Shannon, 1975):

- the model (and/or real world system) is *very complex* with many variables and interacting components, non-linear relationships or random variates;
- the assumptions required by the appropriate analytical model are not sufficiently well satisfied by the real system or the appropriately formulated model has no analytical solution (in other words: a complete *mathematical formulation* of the problem does not exist, is analytically insoluble, or too complex and arduous to solve); and/or
- the possibility of conducting *experiments in the real world* are dangerous, impossible or irreversible (e.g., creating a war to observe its impact on tourism arrivals), too costly (e.g., for processes with long time frames such as the greying of the population time compression may be required), unethical (e.g., the impact of reduced health would require one to actually change these variable in society), and so on.

Simulation-supported analyses can be used to create theories and examine the impact of policy options and/or autonomous developments in society. Impact analyses are executed by comparing simulated changes to a baseline simulation that forecasts the ‘status quo’. This way, simulation is a goal-oriented experiment where the model imitates the system of interest (Merz, 1991). An additional advantage of simulation is its powerful educational and training application because it allows people to see and play with a system (Shannon, 1975).

On the other hand, simulation is not the most efficient and effective answer to each and every modelling problem. The drawbacks of simulation come down to (a.o. based on Shannon (1975)):

- the *development* and *testing* of large and complex simulation models is often expensive and time consuming; this is particularly unfavourable for recreation and tourism organisations that are typically rather small with limited time and resources to develop or apply forecasts (Stynes, 1983);
- there are *no hard rules* about the form of the variables and parameters,

descriptive relationships and constraints, and criterion for judgement of effectiveness (Levine & Lodwick, 1983a); indeed, the construction of simulation models is often referred to as an *art* as much as *science* due to the complexity of the problem;

- simulation is not very precise, and it is difficult to measure the degree of imprecision; there is no such thing as *the* test for the *validity* – rather, the experimenter must conduct a series of tests, including tests for the "face validity" and the use of statistical tests, throughout the process of developing the model in order to build up his or her confidence;
- expert judgement of what is possible is still essential because the simulation system/computer does not initiate or suggest the design of trends, potential new developments and/or opportunities; also, the user, not the system/computer, has to evaluate the results (Shechter & Lucas, 1978);
- simulations are generally valid for one real world system; results do not carry over to other, similar problems (Trick, 1996); and
- the time (and the required CPU-time) to run simulation experiments is usually grossly underestimated (Trick, 1996); on the other hand, using traditional analytical approaches may require even more processing time; in addition, the rapidly increasing processing speed of PC's and the development of advanced computer architectures such as parallel processing and/or supercomputers melts away this counterargument as technology advances.

In conclusion, it can be argued that simulation can be a very attractive tool for experimenting with complex real world situations. However, it is not *the* solution to each and every (complex) problem due to the costs, the imprecision, and the time involved in developing, testing and using a system. In addition, it should be noted that there is no such thing as *the* simulation approach. Rather, like models in general, simulation models can be classified in a number of ways depending on assumptions and solutions. The next section discusses these classification schemes and the accompanying model types.

4.2.2 A classification of simulation models

There are many ways to classify simulation models. Unfortunately, none of these classification schemes is completely satisfactory, although each serves a particular purpose (Shannon, 1975). This section discusses the following schemes:

(a) Deterministic vs. stochastic;
(b) Discrete vs. continuous;
(c) Static vs. dynamic;

(d) Empirically tested vs. non-empirically tested;
(e) Micro vs. macro (or aggregate); and
(f) Language: general vs. simulation.

The following discussion is largely based on: Trick (1996), Merz (1991) and Shannon (1975). Additional sources are mentioned explicitly.

(A) DETERMINISTIC VS. STOCHASTIC

In a simulation model, the mathematically logical relations that describe the behaviour of the real world system can either be deterministic or stochastic. In *deterministic simulation* models, all structural and procedural data are fully determined and the model uniquely determines the simulations. In other words, the output is completely conditioned by a given input, and re-running a simulation will not alter the outcome. In *stochastic* (or *probabilistic*) *simulation* models, on the other hand, some or all of the specified relationships comprise random elements, which implies that the output for a given input is uncertain. In this case, the dependent variable is a random variable that is characterised by a certain probability density function. Drawing from this distribution determines the value of the dependent variable in a certain simulation run. This randomness calls for generating multiple simulation outcomes and the use of statistical measures such as mean value and the standard deviation across the runs to describe the output of the system.

(B) DISCRETE VS. CONTINUOUS

A second aspect that is important in classifying simulation models is the nature of the dependent variable. In *discrete simulation* models, the dependent variable is discrete in nature, and the relationships between dependent and explanatory variables are expressed in terms of conditional probabilities. There are three major ways to approach this type of simulation, including:

- *Event scheduling*, which focuses on the occurrence of events, i.e. anything that changes the state of the system other than the mere passage of time, and that possibly triggers the scheduling of new events; an example in tourism would be the arrival of a visitor in a travel agency and its effect on the (possible) queue or the state of the travel agent (idle becomes busy);
- *Activity scanning*, which focuses on the conditions under which certain activities are executed, and in which – as a consequence - the system is continuously scanned for the satisfaction of these conditions; in the travel agent example, for instance, the activity “start service” is activated when (1) there is a visitor in the queue; and (2) the travel agent is idle;

- *Process oriented modelling*, finally, differs from the previous two approaches by providing tools for the user to define the system; in the travel agent system, for instance, the simulation program would provide (1) a *source* that generates arriving customers; (2) a *queue* to hold waiting visitors; and (3) an *agent* to serve visitors; the modeller, then, needs only provide the parameters such as the interarrival times, the queue discipline and the service times.

For the modeller, the latter approach looks quite different, but from the processor point, the simulation system operates the same as the event scheduling system because events must be scheduled and the system must be updated.

In *continuous simulation* systems, the dependent variables is continuous in nature, and the relationships between the dependent and the explanatory variables are expressed in (interrelated) sets of differential equations rather than probabilities that certain events will take place. Instead of focussing on individual visitors waiting in queues, for instance, the queue itself can be treated as a flow with speed, length and rate(s) of change that are expressed in one or more related equations. The decision whether to select a discrete or a continuous modelling approach depends on the characteristics of the system and the focus of interest of the modeller. Today, hybrid approaches are also available.

(C) STATIC VS. DYNAMIC[12]

A third classification of simulation models is based on the role of time in the system. In *static simulation* models, time does not play an essential role. In the static approach to (tourist) systems, the focus is on the structural (or causal) relationships that are present in the system under consideration. Furthermore, static models assume that the majority of the variables and relationships that constitute the system do not change rapidly. Static models are therefore usually based on cross-sectional data, i.e. information on simulation units at a certain point in time. Experiments, then, are conducted by systematically changing the input-characteristics of the system. This may also include temporal extrapolation to actualise or forecast a sample, which is called "ageing of the sample"[13]. This procedure of "*statically ageing*" a cross-sectional sample is accomplished by (re)weighting the subgroups that make up the sample (under the assumption of invariant relationships over time and generations). These experiments can only tell how much change can be expected, not how fast these changes will take place.

[12] This section is also based on Levine and Lodwick (1983b).

[13] Ageing-procedures are required either to bring the model up to date (frequently, there is a lag of several years between the collection of the data and the application of the model) or to provide estimates for the future (Harding, 1996).

When static models are determined by stochastic relationships between the variables, these models are often called *Monte Carlo* models. Given a set of alternatives with different choice probabilities, for instance, these models determine the outcome by drawing a random number (R) between 0 and 1, and assigning the corresponding alternative[14]. Table 4.1, for instance, displays a(n) (imaginary) choice set of destinations for a summer holiday, including the choice probabilities for the alternatives and the ranges (of the random number) that correspond to these choices. If, for instance, the random number generator returns "R = .3491", a domestic destination will be the outcome of the simulation process. If, in contrast, "R = .9847" is returned, the destination will be outside Europe.

Table 4.1 Example of a choice set for a summer holiday in a Monte Carlo simulation model

Alternative	Probability	Cumulative Probability	Range
Domestic (NL)	.3988	.3988	$R \leq .3988$
Europe (excl. NL)	.5081	.9069	$.3988 < R \leq .9069$
Other	.0931	1.000	$R > .9069$
Total	1.000	1.000	0 - 1

Alternatively, for continuous decision and/or explanatory variables, continuous probability density functions can be used to draw from. In this case, in addition to selecting a model to describe the relationship between the dependent and the explanatory variabele(s), a forecast model for the residuals (actual values less the forecast values) is developed that either takes advantage of the past patterns or incorporates the probability distribution that best captures the distribution of the data. Next, the simulation is run for a large number of trials to obtain forecast of the mean value of the dependent variable and its distribution (Frechtling, 1997).

In contrast to static simulation models, *dynamic simulation* models assume time to be an important factor in a (tourist) system, and the dynamic approach to (tourist) systems focuses particularly on changes over time. These models are often based on time-series (i.e., longitudinal data) and describe (lagged) responses to a changing variable. As a consequence, dynamic models are concerned not only with the magnitude of the impact of these changes, but also very much with the interval of time at which one can expect these changes. An example of (lagged) responses in tourism, for instance, would be the effect of a catastrophe (war, earthquake,

[14] An exemplary *random number generator* assumes a linear distribution between 0 and 1 and returns a value such that (1) no value has a higher probability to occur than another; and (2) two successive values are not correlated.

flooding) on international arrivals: if a catastrophe takes place in a particular region, visitors immediately stay away. Oppositely, however, when the cause of the problem ceases to exist, visitor numbers will not return to the original level immediately, but gradually (if at all).

Another interpretation of the classification "dynamic simulation model" refers to the inclusion of *feedback* mechanism (which may or may not be related to time-related dynamics). The idea behind feedback loops is that the effect of a change in a given state variable X will move through the system and eventually feedback to it again either positively or negatively at a later time. If the variable is part of a *positive loop*, changes in any of the variables around the loop would either lead to a rise or decrease in X, eventually leading to the growth or collapse of the system. *Negative loops*, in contrast, lead to stability because an initial increase if X would eventually lead to a decrease in X and vice versa (Levine & Lodwick, 1983a).

(D) EMPIRICALLY TESTED VS. NON-EMPIRICALLY TESTED

A fourth important feature of (simulation) models is whether the assumptions that comprise its structure are empirically tested or not. If these assumptions are tested, the experiments are referred to as *model- or theory based simulations*. In this context, theory is considered a (system of) assumption(s) that have been verified in (an) empirical setting(s). If, in contrast, the model assumptions are not empirically tested, we speak of *non-empirically tested simulations*.

A simulation system can be tested empirically by comparing simulation results to empirical data. If the system is capable of successfully reproducing empirical data, then there is at least some evidence of the validity of the assumptions (Timmermans & Van der Waerden, 1998). Model assumptions can be tested at (at least) two levels. First, the tenability of the assumptions underlying the components of the simulation system can be tested empirically. If, for instance, a simulation system for tourist destination choices comprises one component that represents the participation choice and one component that describes the conditional destination choice (based on Um & Crompton (1990); see chapter 2), the models describing these two components can be tested on their empirical tenability. Secondly, the entire system can be tested. This way, the assumptions with regard to the relationships between the system's components (including the decomposition of the entire system into its components) are tested.

(E) MICRO VS. MACRO (OR AGGREGATE)

The level of aggregation is also an important distinguishing feature within the context of simulation models. *Macrosimulation* models, for instance, examine relationships between national economic sectors and aggregate variables. These models are therefore also referred to as “aggregate models”. The structure of macrosimulation models is often determined by (a series of complex and interrelated) equations, including, such approaches as regression analysis and gravity-based models. Also, (nested) logit models can be used to develop macrosimulation models. Although macrosimulation models provide important insights into interdependencies between aggregated variables, they cannot be used to assess the distributional impact of (government) policies and socio-demographic changes on individual agents in the economy. In addition, it is very difficult to include mechanisms such as (positive or negative) feedback and threshold-levels in the macro-economic equations. As far back as the 1950s, this prompted Orcutt (1957) in his paper “A New Type of Socio-Economic Systems” to propose the development of simulation models using micro-agents for policy use. Aggregate effects, then, are obtained by combining these individual simulations. Orcutt’s paper originated the field of *microsimulation* models that are directly concerned with the behaviour of microunits (e.g., individuals, households or firms) in different phases of the choice process .

The major difference between the various types of microsimulation models is whether the data are aged “statically” or “dynamically”. As discussed before, the procedure of “*statically ageing*” of a cross-sectional sample is accomplished by (re)weighting the subgroups that make up the sample. In addition, static micro data can be *“uprated”* to account for estimated movements since the time of the survey or anticipated future movements (Harding, 1996).

In *dynamic cross-section microsimulation*, each mirco-unit in the population is aged individually by moving it progressively forward through time (Harding, 1996; Merz, 1991). *Dynamic population models* typically begin with a comprehensive cross-section of an entire sample survey for a particular point in time and age each mirco-unit individually based on survivor probabilities. In contrast, *dynamic cohort models* use exactly the same type of ageing procedures, but usually age only one cohort from birth to death rather than the many cohorts that are represented in an entire population (Harding, 1996). Merz (1991) refers to this procedure as life cycle ageing and to this type of models as *dynamic longitudinal microsimulation*.

In addition to the traditional distinction between static and dynamic types of microsimulation, *agent-based simulation* is one of the new computational techniques (others include cellular automata and genetic algorithms) being pioneered in the ‘complexity sciences’ to examine the behaviour of individual agents (O’Sullivan & Hakley, 2000). An autonomous agent is a computer simulation that is based on concepts from artificial life sciences (Bishop &

Gimblett, 2000). Agents are *autonomous* because each agent has its own *BDI-structure*, i.e. a distinctive set of *Beliefs* (i.e., (perceived) facts), *Desires* (i.e., goals) and *Intentions* (i.e., sequences of (eventual) actions to achieve the goals) that determine the behaviour of the agent without external interventions by, for instance, humans. As such, agent-based simulation models are *rule-based* models because agents have various rules at their disposal that determine how they will respond to the environment and to other agents, and how they will act strategically to achieve their goals. Agent-based models typically examine the interactions between individuals or between individuals and their environment. This focus is accompanied by a one-way notion of emergence of social interactions: the social can emerge from the individual but not the other way around (O'Sullivan & Hakley, 2000).

(F) LANGUAGE: GENERAL VS. SIMULATION

Finally, simulation models can be characterised based on the way they are implemented on the computer. In the context of discrete simulation models, it has already been mentioned that there exist tools or programming languages that provide the basic components (or building blocks) for certain problems. This way, the user only needs to rephrase the problem in terms of the components of the simulation language and provide the parameters of the behaviour under consideration. Examples of these so-called *simulation specific languages* include SIMAN, GPSS, SLAM, SIMULA, DYNAMO, SWARM and DESIRE (to name just a few). Evidently, the major advantage of these simulation languages is that they take most of the work out of creating a simulation system. On the other hand, the application of these tools is limited to their specific domains. GPSS, SIMAN, SLAM and SIMULA, for instance, are designed to simulate dynamic discrete-event systems (Trick, 1996; Shechter & Lucas, 1978). DYNAMO (originally developed by Jack Pugh at MIT), on the other hand, was the first system dynamics simulation language, designed for simulating dynamic continuous problems. SWARM (Minar *et al.*, 1996; Hiebeler, 1994) and DESIRE (Brazier *et al.*, 1997), finally, are examples of languages for agent-based simulations.

Another class of simulation languages is more general and therefore takes away some of the limitations of simulation specific languages. SIMSCRIPT, for instance, is a general language in the sense that things like network optimisation and linear programming can be written in the language. On the other hand it is also a simulation language because it has commands to generate entities and manipulate queues (Trick, 1996).

Finally, simulation systems can be coded in *general programming languages* such as FORTRAN, C++, Pascal and so on. Obviously, the major advantage of these general languages is that there are few limitations to the system apart from the creativity and the programming abilities of the designer. Also, these languages tend to be rather fast, which is an important quality if the system under consideration is very complicated. On the other hand, using a general language to code a simulation system often requires (much) more time to design, code and verify such a system.

4.3 Simulation models in recreation and tourism research

In this section, simulation models in recreation and tourism are reviewed. Given the objectives of this thesis and the focus of the *MERLIN*-system, this inventory is restricted to models that are based on, or aim to assess (the effects of) *choice behaviour of people in (outdoor) leisure settings* and that use a simulation approach because the behavioural system of interest is *too complicated to be handled analytically*. This focus implies that not all models that are referred to as "simulation" or "forecasting" in the original source are included in this review because either the *subject* and/or the *modelling approach* does not match our focus.

The existence or development of simulation models that comply with our focus was explored by scanning the literature and by contacting many experts both in the Netherlands and abroad. Given our focus, however, only 5 simulation systems in recreation and tourism were found, including 3 foreign and 2 Dutch examples. Section 4.3.1 discusses the foreign simulation models and section 4.3.2 discusses the Dutch examples.

4.3.1 Simulation models outside the Netherlands

This section discusses the existing simulation models that were developed outside the Netherlands. Strikingly, the first two examples were developed back in the 1970s and advanced simulation as an "attractive tool" to analysing outdoor recreation planning issues, while the third was developed rather recently. Nothing happened in the mean time and, apparently, the field has not taken up the initial enthusiasm for simulation as a tool for addressing complex planning issues. With regard to the choice behaviour under consideration, all three systems are concerned with choice behaviour at the level of recreation sites to support management decisions. The oldest example, the simulation approach to park planning by Cesario (1975), aims to illustrate the possible advantages of simulation. This model uses a gravity model to describe the generation and allocation of day-trips across various parks in a particular region and includes several assumptions with regard to departure times, queuing behaviour and duration choices to simulate typical patterns of visitor loading in several parks during a particular day. In addition to visitor arrival patterns, the other two examples of simulation models outside the Netherlands aim to capture visitor behaviour within parks and assess the effects of these choices in terms of the number of encounters between (parties of) visitors (which is an important measure for visitor satisfaction within wilderness areas). The remainder of this section discusses these examples in more detail.

A SIMULATION APPROACH TO PARK PLANNING (CESARIO, 1975)

Cesario (1975) was one of the first to advance simulation "as a promising tool for systematic planning analysis" in the field of recreation. In his exploratory paper "A Simulation Approach to Outdoor Recreation Planning" Cesario set forth the principles of simulation modelling and then constructed a simplified model to illustrate the various kinds of park planning issues that can be addressed by use of a simulation approach. To this end, Cesario developed a simulation system with a gravity model as a base. In formula:

$$t_{ij} = \left[g U_i K_i^{(\alpha+1)} \right] \left[\frac{V_j C_{ij}^{\ \beta}}{K_i} \right] \qquad with\ K_i = \sum_{r=1}^{M} V_r C_{ir}^{\ \beta}$$

$$= \left[generation\ O_i \right] \times \left[distribution\ w_{ij} \right]$$

where

t_{ij} : is the number of recreation trips for a specific purpose made from Origin i (O_i; i = 1, .., N) to Destination j (j = 1, .., M) during a specified time period;

g : is a constant of proportionality representing the proportion of total seasonal trips taken on a typical summer Sunday;

U_i : refers to a quantity associated with Origin i and is specified up to a multiplicative constant (Cesario, 1973);

α, β : are behavioural parameters (-1 < α < 0 and β < 0) that, based on the estimation of an operational model, are conveniently rounded off at -0.5 and -1.0 respectively;

V_j's : reflect the differences in the attractiveness of parks and are specified up to a multiplicative constant (Cesario, 1973));

C_{ij} : is the generalised "cost" of travel from Origin i to Destination j, taken to be some function of distance, time and money involved in the trip;

w_{ij} : is the probability that a recreation seeking party from Origin i chosen at random will travel to Destination park j.

Cesario used the General Purpose Systems Simulator (GPSS) because it is particularly well-suited for modelling the movement of items (people in this case) over networks (the recreation system consisting of N origins, M destination parks, and a set of at least NM travel links). The model was developed with particular reference to an outdoor recreation system in north-eastern Pennsylvania ($N \times M$ = 12×5), and one unit of simulation time equals one minute of real time of a typical summer Sunday between 8 A.M. to 8 P.M. The mean time between trips from a particular Origin county i to a particular Destination park j (intergeneration times) are assumed arbitrarily to be exponentially distributed, and to vary with the Origin county i and the time period of the day. As each new party is GENERATEd[15,16] the distribution of vector $w_i = (w_{i1}, w_{i2}, .., w_{iM})$ is sampled, and the party is assigned to a park according to the random number that is obtained. Next, the GENERATEd party is ADVANCEd over the highways where it must queue to pay the entrance fee. If the queue size is 20 or less, the party waits; if it is greater than 20, the party elects either to return home or to visit another park, the decision being based primarily on the relative travel times between returning home (the trip is TERMINATEd) and visiting the next nearest park. Finally, the model also incorporates visitor lengths of stay. More specifically it assumes the expected length of stay to vary, to be exponentially distributed, to vary across the parks and to be longer for early arrivals. At the completion of a stay the trip is TERMINATEd.

Using the above assumptions, Cesario (1975) performed several experiments, including the introduction of alternative locations for a new park and the implications of raising the prices at selected sites (by increasing C_{ij}). The output of the experiments included (a) visitor loading; (b) capacity utilisation; and (c) queue length distributions. Finally, Cesario (1975) tentatively concludes that simulation could indeed be useful for analysing recreation planning problems.

THE WILDERNESS USE SIMULATION MODEL (WUSM) (SHECHTER & LUCAS, 1978; SMITH & KRUTILLA, 1976).

The second foreign example is the Wilderness Use Simulation Model (WUSM). A prototype version of the WUSM was developed by Smith and Krutilla (1976) to

[15] The capitalised words such as GENERATE, ADVANCE, TERMINATE, etc. represent the names of “blocks” used in GPSS to accomplish certain functions .

[16] A GENERATE block is included to represent visitors from all excluded counties.

study encounters between parties of visitors under varying conditions of wilderness usage in the Spanish Peaks Primitive Area in Montana. Later modifications and pilot tests by, amongst others, Shechter and Lucas produced a more complex and flexible second-generation WUSM. Formally, the WUSM assumes a wilderness to be a type of processing operation in which parties arrive at the area according to some specified temporal pattern. Next, parties select routes and execute them. For this purpose, the WUSM requires the following input (supplied by the user):

(1) Area characteristics, i.e. the route network that includes 4 physical features (see Figure 4.1): trailheads, individual trail segments, campsites, and travel routes (i.e. a series of trail segments and campsites that are connected in a sequence);

(2) User characteristics, where the basic unit is the party characterised by mode of travel (e.g., hiking vs. riding, paddle vs. canoes), size (e.g. large, medium, small), and arrival patterns that are represented by probability functions (based on four nested factors: the total level of use for the entire season, the weekly distribution within a season, the daily distribution within a given week and the hourly distribution for a given day);

(3) Route-user interactions (travel patterns) that comprise information on transit times for trail segments and campsites (composed of a base transit time (in minutes) for each trail segment, and a percentage modifier for the particular type of party), and on the likelihood of selecting particular trailheads, segments and campsites which may depend on, temporal factors (period of day, type of week/day) and on characteristics of the travel party; the WUSM assumes travel parties to make up their travel patterns at the trailhead; and

(4) Run length specification, i.e., the number of times the simulation run should be replicated.

Given the above information (i.e. databases), the WUSM passes simulated parties through the model and records encounters (i.e. user-user interactions). These encounters constitute the decision variable and include the following categories:

(1) Camp encounters in which parties use the same camping area overnight;

(2) Meeting encounters in which parties approach each other from opposite directions on a trail segment and pass;

(3) Overtaking encounters in which one party on a trail segment overtakes and passes another party; and

(4) Visual encounters in which one party – whether on trail or in camp – sights another one travelling on a different trail segment or campsite.

Depending on the version (I, II or III), the WUSM's output includes 21-29 databased tables (i.e. tables of input data that allow the user to verify that simulated parties are entering the model as intended), 9-11 output tables that show use (segment-oriented summaries) and encounter data, and (version III only) detailed day-by-day and party-by-party records on magnetic tape.

Like Cesario's (1975) simulation approach to park planning, the WUSM is coded in the General Purpose Simulation System (GPSS) computer language (version GPSS V) because this language is specifically suitable to deal with scheduling problems in discrete-event systems (i.e., time-dependent systems in which changes in state occur at discrete points in time). The output analysis tools are written in FORTRAN.

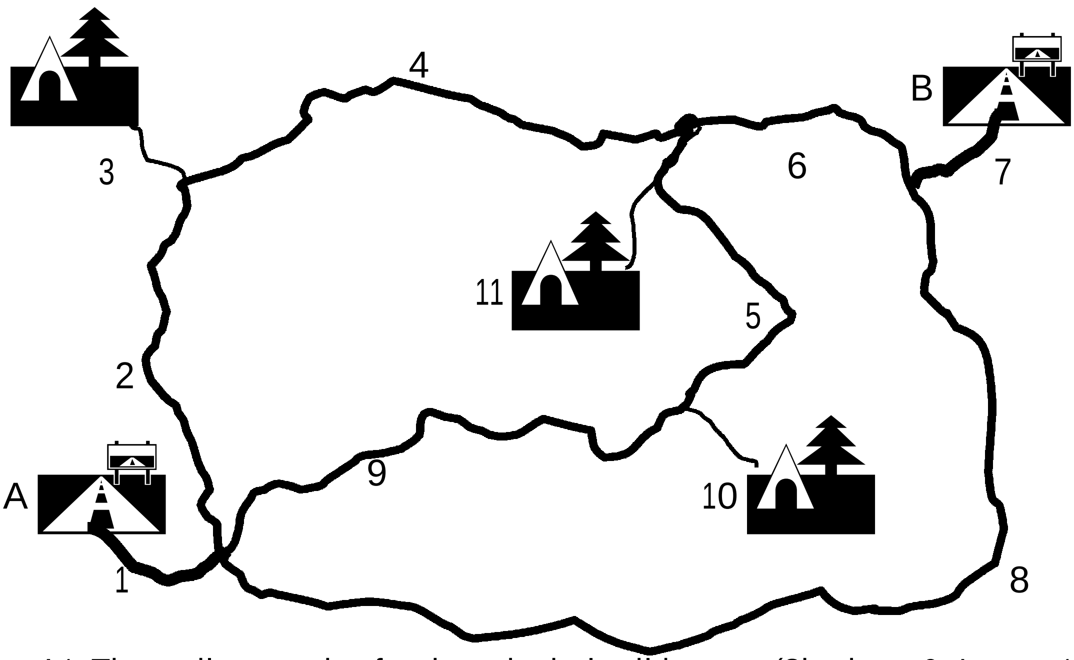


Figure 4.1 The trail network of a hypothetical wilderness (Shechter & Lucas, 1978), including two trailheads (A & B), eight trail segments (1,2, 4, 5, 6, 7, 8 & 9) and three campsites (3, 10 & 11);

RBSIM – RECREATION BEHAVIOR SIMULATOR (GIMBLETT & ITAMI, 2000; BISHOP & GIMBLETT, 2000; GIMBLETT, 1996; KOLSTÉ, 1997)

The Recreation Behavior Simulator (RBSim) by Randy Gimblett (University of Arizona, USA) and Bob Itami (Director of Digital Land Systems Research, Parkville, Victoria, Australia) is an *agent-based* model that simulates the behaviour of visitors of high use natural environments. Although still experimental at the moment, RBSim demonstrates the potential progress that is at hand when two technologies – Geographic Information Systems and Autonomous Human Agents – are combined to explore the complex interactions between humans and the environment. In effect, RBSim is a modern, more technologically advanced successor to the Wilderness Use Simulation Model, because the principal aim of the system is to study the number and type of interaction visitors will have within each group and between groups. Figure 4.2 shows the possible configuration of a Spatial Decision Support System (SDSS) that could be developed based on RBSim. In this SDSS management options are processed through the system to generate visitor satisfaction levels and environmental impact levels. The sustainability of the proposal can thus be assessed and adjusted. Additional GIS-data sources, agent types or environmental impact models can be added as appropriate and available.

There are two stages in the process of establishing behavioural rules for agents in the recreational context. First, there needs to be an assessment of agent types, i.e. visitor types based on, for instance, mode of travel, purpose of visit or preference for crowds or isolation. Next, specific responses to possible environmental conditions (physical or perceptual) should be determined[17]. In a case study in Broken Arrow Canyon, Arizona, Gimblett (1998) distinguished between *hiking agents* (subdivided into social experience-oriented (*gregarious*) and landscape experience-oriented (*isolationists*)), *mountain-bike riding agents* (again subdivided into gregarious and isolationist agents) and *jeep-tour agents*. The behavioural rules of the agent types in this study were based on 'reasonable' suppositions (partially supported by/based on empirical evidence from interviews and observations) and defined in terms of (1) range of movement, where hikers, bikers and jeeps each have their own fixed trials modelled as linked available cells in a raster database; (2) speed of movement (dependent on the agent type and the steepness and direction of slopes); and (3) variations in speed dependent on age, recent energy loss or gain (by rest), and the tendency to join with others or to avoid others (both physically and visually, where the possibility to view (and avoid) other agents depends on terrain conditions).

The model runs under the RBSim software that was developed using routines drawn form the raster grid software package SAGE (SAGE has recently been converted to an object oriented programming language C++, which eases the communication between the GIS-database and the agent languages). The human dimension is implemented by agents developed from the two types of multi-agent systems: the *intentional* agent programming language dMars (Distributed Multi-Agent Reasoning System) developed at the Australian Artificial Intelligence Institute AAII, and the SWARM programming language that incorporates

[17] Surveys are required to determine both visitor types and conditional responses. These surveys can be based on field interviews, questionnaires or observations, or on people experiencing a virtual reality (Bishop & Gimblett, 2000).

responsive agents. This way, RBSim incorporates information on the information systems of individual agents as well as the joint behaviour of tourists.

The user determines how many agents of each type are to be used in the simulation process, what age groups they fall into, and the time intervals between them setting out (the latter parameter can be randomised within constraints). Also, the user establishes the total duration of the simulation and renders active or inactive the rules that the agent types use for testing purposes (e.g., all agents stop at all landscape features; hikers and bikers will not stop at features if more than 5 other agents are present). Based on summary statistics RBSim can map or graph various indicators of conflicts or contentment, including, for instance, the number of encounters between two agent types or the level of visitor satisfaction of a particular agent type (based on the surveyed objectives of the agent type). Using these outputs, RBSim can assess different trail configurations, the contribution of permitted visitor numbers or activities (and different behavioural rules) to the overall tourist satisfaction.

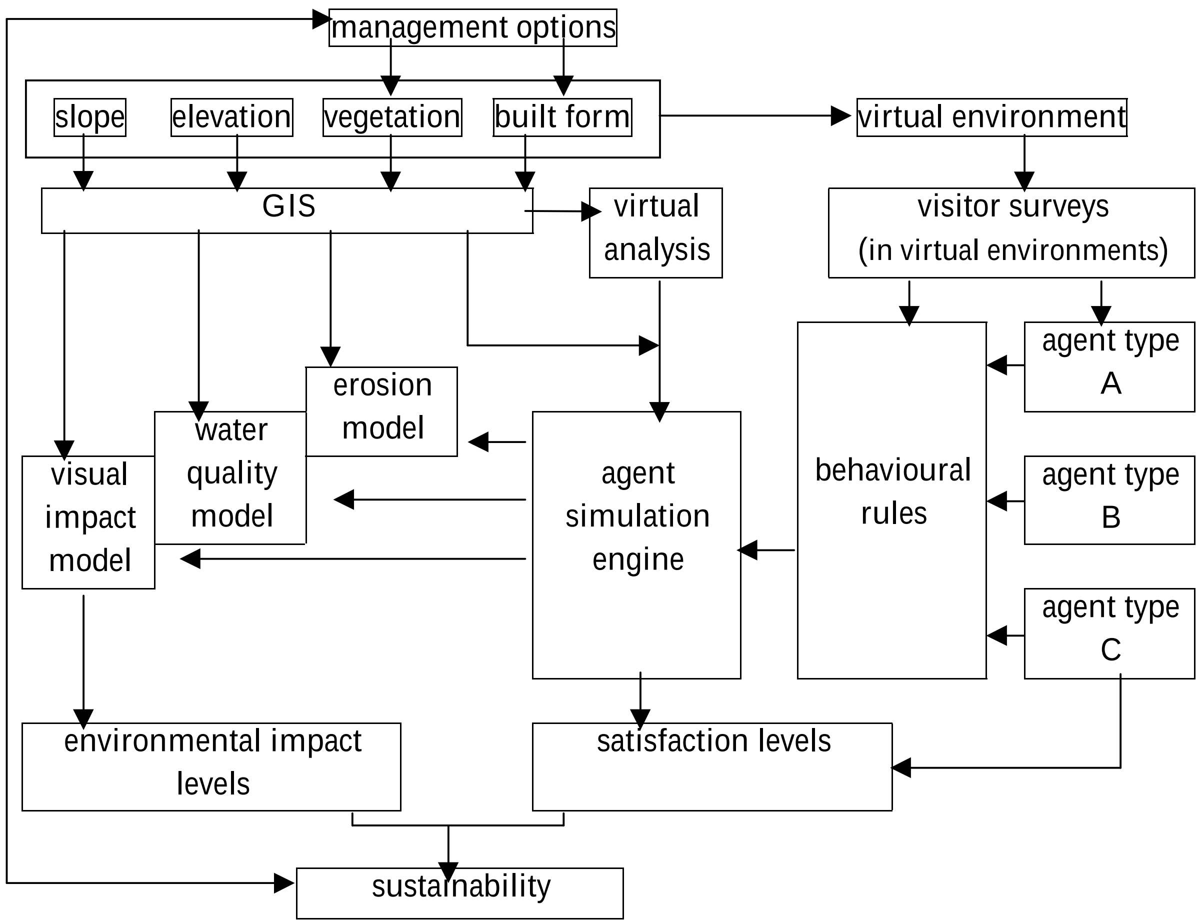


Figure 4.2 The concept of integrating GIS and autonomous agent modelling for sustainable natural area recreation management (Bishop & Gimblett, 2000)

4.3.2 Simulation models in the Netherlands

In contrast to the foreign examples, the Dutch simulation models have both been developed in the 1990s and are concerned with choice behaviour at the national level rather than the site or the regional level. Also, in addition to modelling the demand for various tourist and recreation activities, the focus is often on the economic consequences of these choice processes. Hence, these models are typically suited for policy makers and planners at the national level (although some elements of the models may be interesting for regional and local planners as well). The models were both developed under the authority of the Dutch Ministry of Economic Affairs as part of a comprehensive macro-economic forecasting and decision support system for tourism in the Netherlands. This comprehensive system comprises the day-trips and holidays of both Dutchmen (and in the case of holidays also foreigners) in the Netherlands. The remainder of this section discusses the existing simulation systems in the Netherlands in more detail.

SEO'S "DAY-TRIP MODEL" (BROUWER *ET AL.*, 1995)

The Day-Trip Model (DTM) developed by the Foundation for Economic Research (SEO) of the University of Amsterdam aims to forecast the annual amount of money that will be spend on day-trips by the Dutch population. The DTM is a microsimulation model that conceptualises the choice process for day-trips as having three components. These components are presented in Figure 4.3 that also shows the level of analysis of each component, the modelling approach that is used, and the possible effects of this component on the simulation results. The DTM uses the 1990/'91 day-trip survey (CBS, 1992) to calibrate the three components. The explanatory variables for the various components include socio-economics and -demographics (age, income, gender, social situation, education, working situation), geographic information (number of inhabitants of the destination, the residential region of the tourist), weather conditions and the type of day (week/-end, holiday period).

The first component represents the number of day-trips using an ordered logit model. This model only uses explanatory variables at the level of the individual (including an individual correction factor for the possible type of day-trip activities; see hereafter) and predicts the respective probabilities for no, one, two, .., etc. day-trips. The model also includes an intercept for each possible number of day-trips. The maximum number of day-trips day-trips is 32, where 10 numbers are missing (based on the data). The model thus requires 22 intercepts. In the simulation, the expected number of day-trips *E(m)* a person will make is given by:

$$E(m) = \sum_{m=0}^{32} P(number\ of\ \ day - trips = m) * m$$

where *m* represents a particular number of day-trips, and *P(number of day-trips = m)* is the probability for *m* day-trips as predicted by the ordered logit model.

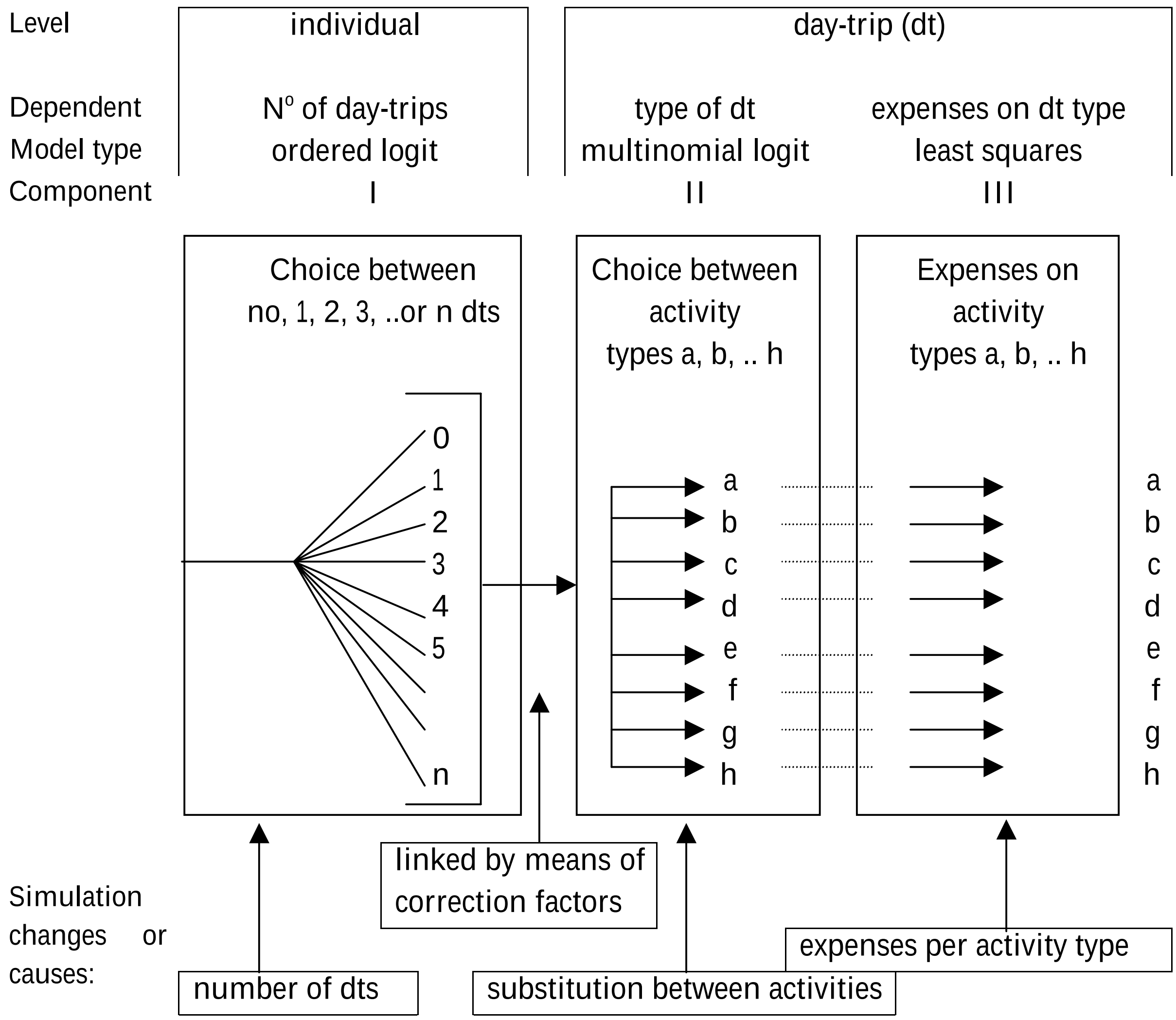


Figure 4.3 Schematic representation of SEO's "*Day-trip Model*" (Brouwer *et al*., 1995)

The second component of the DTM determines for each day-trip the type of activities that will be pursued, including outdoor recreation, sports activities, visiting places/objects of interest, attending shows, shopping, hobby's, visiting amusement parks and visiting catering establishments (where "visiting catering establishments" is the reference alternative, and all (personal and day-trip) variables are entered as alternative specific parameters). The DTM assumes that there exists a *relationship between day-trip frequencies and activity types* (e.g., the sports and hobby's are high-frequent activities, whereas visiting amusement parks are low-frequent activities) and introduces correction factors to represent this relationship. In practice, the Multinomial Logit Model that represents the second component is calibrated first (representing the probability that a particular activity type will be selected conditional on the individual's day-trip frequency). Using these model parameters, the following correction factor is calibrated for each individual:

$$\ln(C) = \frac{1}{T_k} \sum_{i-1}^{T_k} \left\{ \ln[1 + \sum_{j}^{S-1} \exp(X\beta_j)] \right\}$$

where ln(C) is correction factor for individual *k*, T_k is the number of day-trips of individual *k*, *i* represents the i^{th}-day-trip of individual *k*, *j* represents the day-trip type (*j* = 1, 2, 3, .. S), *X* is the vector of explanatory variables for the choice of activity type *j*, and β_j is the vector of parameters for these explanatory variables. The correction factor, then, is used in the ordered logit model as an explanatory variable.

The third and final component represents the amount of money that will be spend on each activity type using eight OLS-regression models (one for each activity type) that are independent from each other and from the other two components.

Finally, given these components, the amount of money that will be spend on each day-trip type is broken down into entrance fees, consumption and travel expenses (on car/other mode) using a (fixed) matrix that represents the share of these types of expenses for each day-trip type (where, overall, consumption dominate the total expenses with an average share of 63%). The DTM is used to assess the effects of changes in the social-economic and –demographic composition of the population on day-trips and on expenditures on these trips.

TOERMODEL (BORGERS & TIMMERMANS, 1996; SEWRAJSING, 1996; VELTHUIJSEN & VERHAGEN, 1994; VAN DIJK *ET AL.*, 1991)

Toermodel is a macro-economic simulation model of both domestic and incoming holidays based on time series data on holiday expenditures. According to Van Dijk *et al*. (1991), Toermodel is designed to perform both scenario-evaluations and forecasting. The forecasting horizon is 1-5 years (short and mid-term).

Toermodel was initially developed by SEO in commission by the Ministry of Economic Affairs (Velthuijsen & Verhagen, 1994; Van Dijk *et al*., 1991). Later, the exploitation and further development of Toermodel was commissioned to TNO-

INRO (Sewrajsing, 1996). Finally, several submodels were updated by the European Institute of Retailing and Services Studies (EIRASS) of the Eindhoven University of Technology (Borgers & Timmermans 1996). A schematic representation of the various submodels of Toermodel is presented in Figure 4.4.

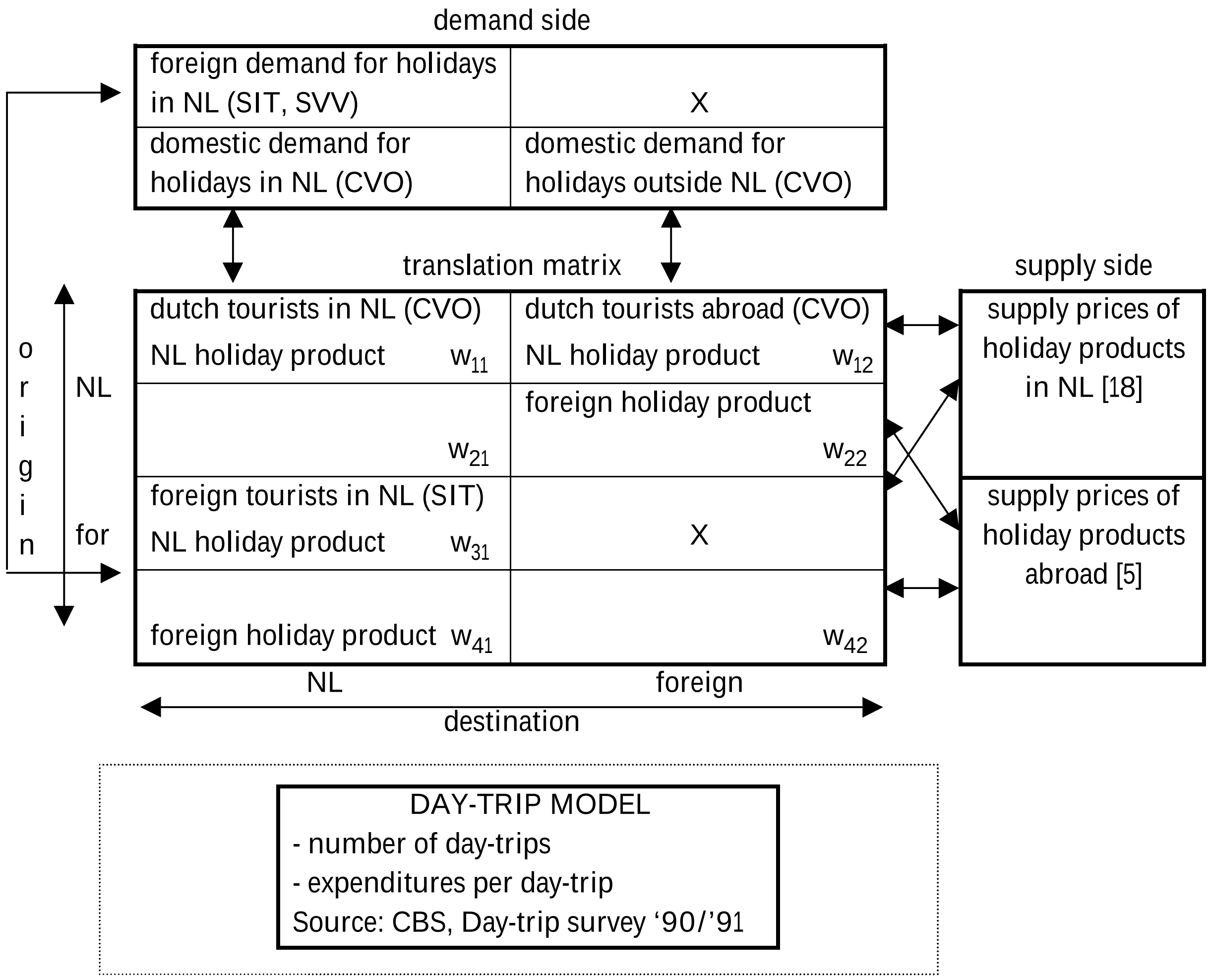


Figure 4.4 Schematic representation of the relations between submodels of Toermodel (w_{ij} indicate submatrices of weights in the translation matrix) (Sewrajsing, 1996)

Basically, Toermodel consists of a demand and a supply side and a translation matrix for holidays. In addition a separate (but related) day-trip model is available (the latter model has been discussed previously (Brouwer *et al.*, 1995)). The demand side of the market comprises a part for incoming holidays and a module for the holiday behaviour for Dutch tourists. The module for incoming holidays, in its turn, first comprises a participation model that, based on time-series models, specifies the probability that an inhabitant of a foreign country will visit the Netherlands for a holiday, business visit of family meeting. Forecasts are based on variables like changes in income, inflation and exchange rate (Sewrajsing, 1996; Van Dijk *et al.*, 1991). Next, a series of logistic regression models explain the choice of the type of holiday in terms of accommodation (4 alternatives) and duration choices (2 alternatives). Finally, the amount of expenditures given the type of holiday is modelled by means of an ordinary regression analysis (see Figure 4.5). Characteristics of the visitors and their travelling parties are used as explanatory variables. The data used for the modelling system were obtained from the survey on incoming tourists (SIT) by Statistics Netherlands.

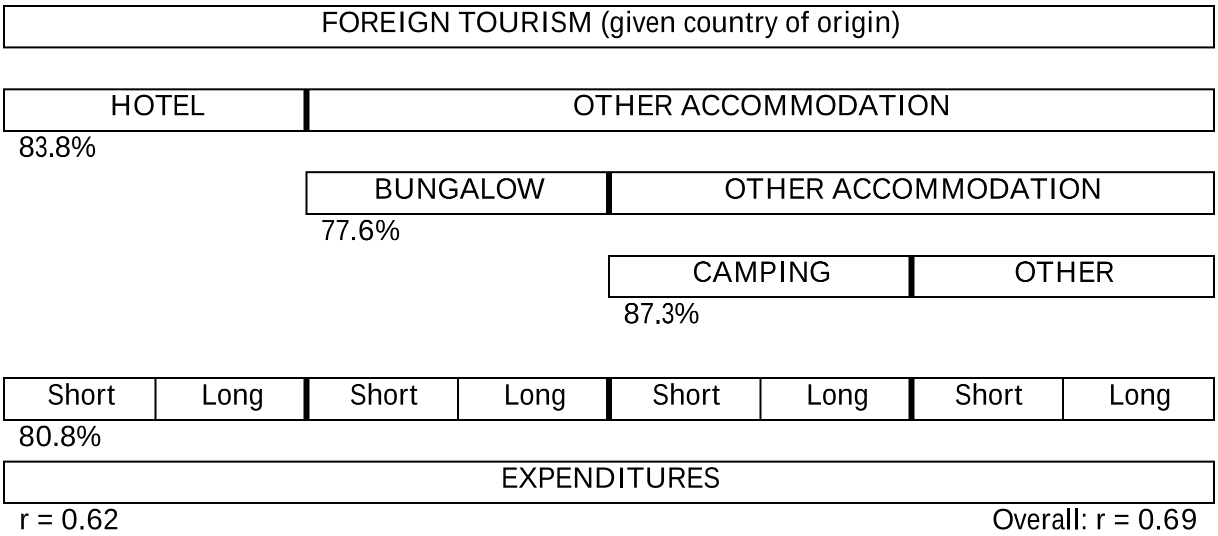


Figure 4.5 Modelling system for foreign tourists in Toermodel (including the percentages of correctly predicted by the logistic regression analysis and the r for the ordinary regression analysis)

The modelling system for Dutch tourists consists of three modules (see Figure 4.6). First, a series of logistic regression analyses determine the probabilities that an inhabitant of the Netherlands will be involved in at least one, two, or three (or more) holidays. These models are estimated using characteristics of the respondent and the corresponding household as explanatory variables. The second module computes the probability that a holiday will be taken in the Netherlands or abroad again using a logistic regression model and characteristics of the respondent and the corresponding household as explanatory variables. In addition, dummy variables are included to indicate whether the holiday under consideration was a first, second or third (or higher) holiday. With regard to the Dutch destinations, the Netherlands are broken down into 8 regions (beach; water sports; forests, hills, central; forests, hills, north-east; forests, hills, south; luxury bungalow parks; other areas). The third and final module computes the amount of money that will be spent by the tourist using ordinary regression analyses. Model estimates are based on the 1994/1995 Continuous Vacation Survey databases (CVO); respondents were weighted to represent the Dutch population.

The supply side of the market is regarded as competitive and demand determined, and it is assumed that the larger part of firms within the tourist market are price-takers: the firms produce and supply at prices that as a rule are set equal to marginal cost. As a consequence, the supply side of Toermodel consists of price equations of various sectors (instead of equations describing quantities supplied).

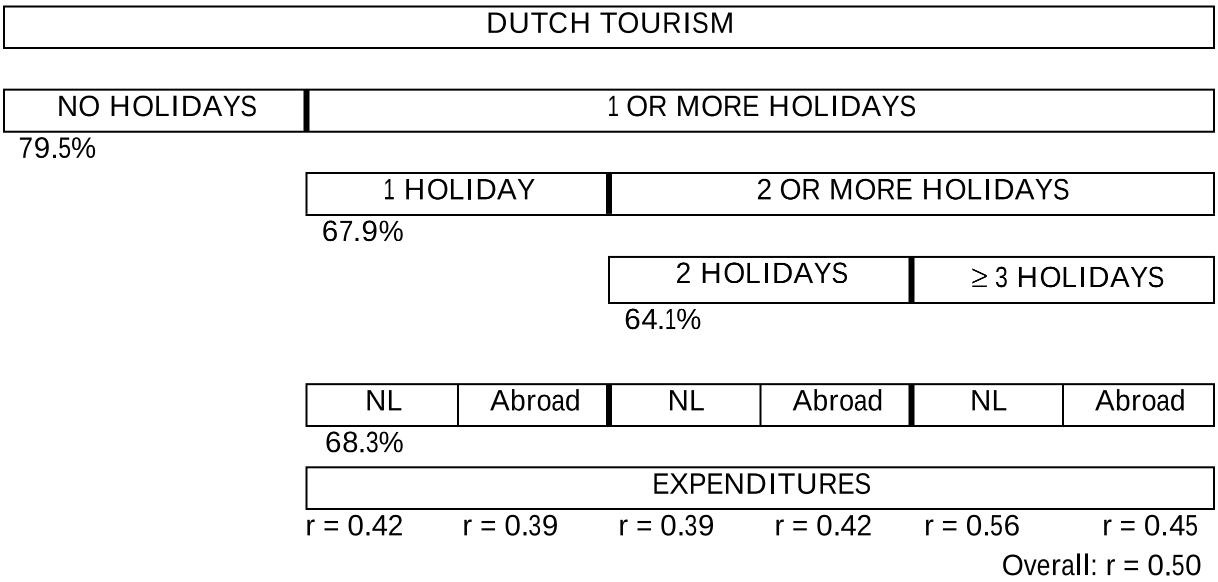


Figure 4.6 Modelling system for Dutch tourists in Toermodel (including the percentages of correctly predicted by the logistic regression analysis and the r for the ordinary regression analysis)

The following sectors are included: food and accommodation industries (8 subsectors), transport (5 subsectors), entertainment (3 subsectors), retail trade (shops) and intermediaries (tour operators). Exogenous factors, as well as variables expressing the effects of policy instruments such as taxes, social security premiums, duties and subsidies are among the explanatory variables. The rough general specification for each single industry reads (Van Dijk *et al*., 1991, p. 16):

$$\Delta \log P = (\alpha_1 \Delta \log PL + \alpha_2 \Delta PI + \alpha_3 \Delta \log PO) + \Delta \log M$$

with $\Delta \log M = \beta_1 \Delta \log AT + \beta_2 \Delta OC + \beta_3 \Delta \log TT + \beta_4 \Delta N$

where: *P* = price of deliveries
PL = unit labour costs
PI = unit cost of purchases
PO = unit cost of other inputs, excluding capital costs
M = profit margin
AT = average turnover of a firm
OC = occupancy rates
TT = total turnover (of all other firms together)
N = number of firms in the sector

The translation matrix, finally, is explained as follows (Van Dijk *et al.*, 1991, p. 55): "Suppose we know the evolution of prices in the tourist industries. By multiplying the vector of price indices of the industries (classification 1) with the translation matrix, we can calculate the (vector of) prices of the holiday types as the tourist sees it (classification 2). Alternatively, if we know for instance the demand for the various holiday types, we can calculate the consequences for turnover per industry, again by multiplication with the translation matrix. The elements of the matrix, which we will designate as w_{ij} represent fractions (in NLG) of holiday type *j* supplied by industry *i*. Or, less formally, the matrix contains the weights of the specific industries in the types of holidays." The weights are based on a large number of sources for the period between 1976 and 1989, including, for instance, hotel and bungalow guides, monthly statistics with regard to socio-economic issues and prices, sector databases, and so on.

4.3.3 Conclusion and discussion

With regard to the existing simulation models both in and outside the Netherlands, only 5 examples could be identified. Although this may, in part, be attributed to the rather strict definition of simulation and/or the focus with regard to the subject, this observation does not deviate that much from other disciplines. Despite the early launch and distinct advantages of the simulation approach, in many ways the field is still relatively new and the use of simulation models, *micro*simulation in particular, have not become routine. In part, this is due to the demanding computing requirements of some of the microsimulation models. More importantly, while computing obstacles are rapidly diminishing, a more fundamental problem is the lack of good data – both about the current characteristics of the population and about the forces driving people's behaviour (Harding, 1996).

A classification (as far as possible) of the simulation models in this review is presented in Table 4.2. In general, it can be concluded that the majority of systems is not empirically tested. Or, at least, these tests are not reported in the literature. This may be related to the fact that it is very time and money consuming to collect the required data.

Table 4.2 Classification of existing simulation models in recreation and tourism

	Deterministic or Stochastic	Discrete or Continuous	Static or Dynamic	Empirically tested?	Micro or macro	General or Simulation Language
Foreign models:						
Park planning model	S	D	D	no	Mi	GPSS (S)
WUSM	S	D	D	yes	Mi	GPSS (S)
RBSim	S	D	D	no?	Mi	SAGE+dMars (S)
Dutch models:						
Day-trip model	rule (S?)	mix	S	?	Mi	unknown (G)
Toermodel	S	mix	S	yes	mix	unknown (G)

With regard to the identified simulation models there are some striking differences between the foreign and the Dutch examples. First, the foreign examples are, without exception, dynamic discrete-event systems, whereas the Dutch are static in nature and have both discrete and continuous components. In addition, the foreign systems use (at least in part) simulation languages while the Dutch systems (although not mentioned explicitly) are written in general purpose languages. Finally, the foreign systems simulate individual or group choice processes (microsimulation), whereas the Dutch systems mainly focuses on macro-economic processes. This is related to the fact that the subjects of the Dutch and foreign systems differ significantly. While the foreign examples focus on the demand for and/or behaviour at the site level, the Dutch examples are concerned with demand for recreation and tourism at the national level. Finally, in addition to modelling the demand for various tourist and recreation activities, the Dutch focus often also includes the economic consequences of these choice processes. It should be emphasised that this is not representative of all (quantitative) models in the field of recreation and tourism in the Netherlands because there are many Dutch studies and models that focus on other aspects of tourist and recreation choice behaviour. These models, however, do not qualify as simulation models.

4.4 Conclusion and discussion

This chapter provided the final conceptual building blocks for the development of the *MERCIN*-system. First, simulation was introduced as an approach to analysing complex problems that cannot be analysed analytically. In essence, the *simulation approaches* build representations of complex real-world situations by synthesising from these relatively simple parts. The principles underlying simulation approaches were reviewed by discussing the (dis)advantages of simulation, the circumstances under which simulation offers an attractive tool for solving or analysing problems and several schemes to classify simulation models. Finally, the application of the simulation in the field of recreation and tourism was reviewed.

5 The *MERLIN*-System

This chapter completes the (description of the) development of the *MERLIN*-system. Based on the conceptual framework presented in chapter 2, this chapter formalises the modelling problem of the *MERLIN*-system. Next, the system's architecture is discussed.

5.1 Introduction

The *MERLIN*-system is a microsimulation model of tourist trip patterns that is based on the underlying decision-making process. The previous chapters have outlined the underlying conceptual considerations, the definitions and the scope of the system and the modelling considerations. The final decision comprises the identification of the components that perform specialised functions in the system. To this purpose, this chapter first formalises the modelling problem. Subsequently, the system's architecture is outlined, including a discussion of the system components and their tasks.

5.2 The formal modelling problem

The *MERLIN*-system comprises the most important facets of the travel decision that are (often) taken long before departure and/or that are relevant to policy makers and planners at the national level. Based on the conceptual framework, these choice facets typically include the travel party, the timing of the trip, the destination, the accommodation, the transport mode, the date of departure and the expenditures. Also, a distinction is made between main and additional tourist trips, where more important trips are considered prior to less important trips. The remaining parts of this section discuss and formalise this process.

Let P_i represent the set of the personal and household characteristics, and S_i the set of attributes and features representing the system and the institutional context (see chapter 2, section 2.4) of an individual i (i = 1, .. I). Furthermore, let the *tourist trip program* ($\wp_i$) be defined as the annual set of tourist trips $\mathbf{T}_{ih}$, (where h is an index that identifies the trip; h = 1, .. , H) that individual i pursues during a particular year given his or her personal and household characteristics and the system and the institutional context. In formula:

$$\wp_i = \{\mathbf{T}_{ih} | P_i; S_i\} \;\; \forall \;\; i, h$$

If individual i does not make any tourist trips in a particular year (H = 0), the tourist trip program is an empty list $\wp_i = \{\varnothing \mid P_i; S_i\}$. Once an individual's tourist trip program has been generated, the next step is to characterise each tourist trip ($\mathbf{T}_{ih}$) by its profile, including the interrelated choices of duration or length of the trip in days *l* which is an element of the choice set for duration ($l \in §_L$), the travel party or companions *c* ($c \in §_C$), the timing *s* ($s \in §_S$), the destination *d* ($d \in §_D$), the accommodation *a* ($a \in §_A$), the travel mode *m* ($m \in §_M$), the date of departure *w* ($w \in §_W$) and the expenditures per trip € (€ > 0); (the choice sets for these facets are detailed in the chapters that detail the empirical building blocks of MERLIN). We refer to this process as *trip profiling*. Once all trips in the tourist's trip program have been profiled, the list of profiled trips is referred to a *planned tourist trip pattern*. During the implementation, the planned trip pattern is implemented and adjusted to account for unforeseen circumstances, ultimately resulting in the *(observed) tourist trip pattern* ($\wp_i$*).

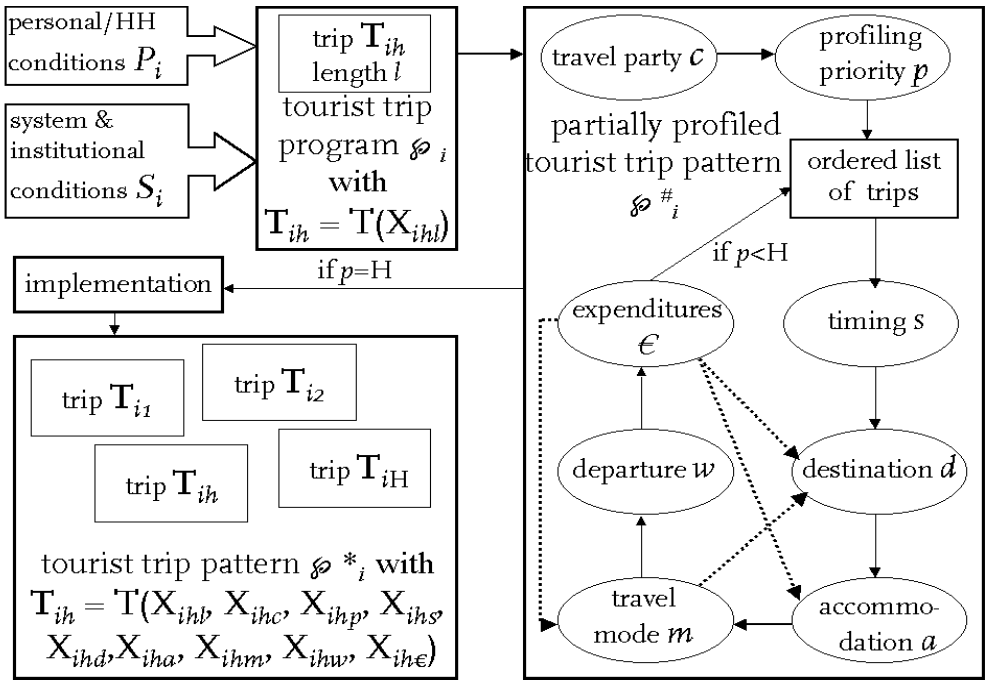


Figure 5.1 The stages of the tourist decision-making process in the MERLIN-system

MERLIN assumes a sequential trip profiling process. Evidently, the order in which *profile choices* are made will affect the results because previous choices frame the condition space of subsequent choices. The order of the choice facets within the trip profile was established based on the most likely direction of the relationships between the various choice facets (see section 2.4 for details).

However, some relations, such as the relationship between destination and transport mode choices may be bi-directional, where, in some cases the destination, and in other cases the transport mode is contemplated first. Also, some choices may be taken simultaneously. This problem is inherent to assuming a sequential decision-making structure. MERLIN partially avoids these problems by including conditions that refer to decisions yet to come (this aspect is worked out in particular with regard to expenditures (expenditures on trips with higher profiling priorities are included in some cases as conditions for decisions yet to come) and the relation between destination and mode choices; see chapters 10 and the dotted arrow between destination and transport mode in Figure 5.1).

The order in which the *trips* are profiled will also affect the outcome of the profiling process because decisions on other trips may affect the profiling decisions on trips yet to come. The order of trips in the tourist's trip program represents the order in which these trips will be considered in the scheduling process by individual *i* rather than the order in which they will be undertaken during a year. This order is determined by the *profiling priority p* of a tourist trip (*p* = 1, 2, .. , H; where '*p* = 1' indicates the most important trip and '*p* = H' indicates the trip with the lowest profiling priority). To reduce the complexity of the model system, the profiling priority of each trip within an individual's tourist trip program is first based on the *duration* (the longer the trip, the higher the priority) and second, if the duration is identical, on the *travel party*. As a consequence, the priority of a trip within the profiling process is determined following the choice of travel party and it (1) distinguishes between main holidays that are of often longer, more distant and/or more expensive (Dirven *et al*., 1998), and extra holidays; (2) reflects the idea that longer trips often have a more extensive planning horizon; and (3) establishes the commitment to other people and their tourist trip patterns. With regard to the latter commitment and the choice set for the travel parties ($§_C$) that will be identified in chapter 6, *trips with schoolchildren* have the highest priority because these travelling companions have fixed (school and holiday) agenda's; trips for *single persons*, on the other hand, have lowest-but-one priority because the traveller does not have to anticipate other people's schedules (trips where the travel party is *unknown*, have the lowest priority). For similar reasons, trips with *children between 0-5 and 15-19 years* old, travel parties with *adults only* (> 20 year) and with travel parties of *9 or more people*, respectively, are in between.

Once a profiling decision with regard to a particular trip has been taken, it is added to the *partially profiled tourist trip pattern* $\wp^{\#}_i$ for individual *i*. In the profiling process, the characteristics of the partially profiled trip pattern, including previous decisions regarding facets of the trip under consideration as well as those on trips with higher profiling priorities, are included as conditions in choices yet to come.

Given the above assumptions and definitions, the modelling problem of the MERLIN-system can formally be expressed as predicting the following tourist trip pattern $\wp^*_i$ for each individual i (i = 1, .., I):

$$\wp^*_i = \{\mathbf{T}_{ih} \mid P_i; S_i\} \quad \forall\, i, h \qquad \text{with}$$

$$\mathbf{T}_{ih} = T(X_{ihl}, X_{ihc}, X_{ihp}, X_{ihs}, X_{ihd}, X_{iha}, X_{ihm}, X_{ihw}, X_{ih€})$$

where X_{ihn} represents the n-th ($n = l, c, p, s, d, a, m, w,$ €) facet of h-th trip $\mathbf{T}_{ih}$ in the trip program $\wp_i$ of individual i; these facets include the duration (length of the trip in days) l is an element of the choice set for durations $§_L$ ($l \in §_L$), the travel party or companions c ($c \in §_C$), the profiling priority p (p = 1, 2, .. , H), the timing (season) s ($s \in §_S$), the destination d ($d \in §_D$), the accommodation a ($a \in §_A$), the travel mode m ($m \in §_M$), the date of departure (part of the week) w ($w \in §_W$), and the expenditures € (€ > 0); and this whole choice process is completed given the sets of personal and household (P_i) and system and institutional (S_i) conditions. The conditions originating from partially profiled trip pattern $\wp^{\#}_i$ are not included in the formal modelling problem because they are endogenous, i.e., temporary conditions that only operate during the decision-making process.

To close, it is noted that the choice sets for each phase of the decision-making process (i.e. the choice alternatives) will be presented and substantiated in the chapters 7 through 11 inclusive that discuss the empirical building blocks of the MERLIN-system. Table A1.2 in Appendix 1 provides an overview.

5.3 The architecture of the MERLIN-system

Concluding the development of the MERLIN-system, this final section outlines the system's architecture. Schematically, the components of the MERLIN-system are presented in Figure 5.2.

The main functionality of the system, i.e. the simulation of the tourist choices of a given population based on a set of decision models and including the storage of simulation results in various formats, is represented in component B, the *Simulation Process*. The population under consideration is generated by uprating and static ageing a standard set of simulation units (the reference population) and by defining the system and institutional context based on a user-defined “policy”[18]. This process is represented in component A, the *Population Generator*. Thirdly, the stored simulation results and the population under consideration can be inspected by cross-tabulating the simulation results and/or the socio-economic characteristics of the population. This process is represented in component C, the *Table Generator*. Finally, the whole system can be influenced by the user via the *User Interface*. The following subsections will subsequently detail the structure and specialised tasks of each component.

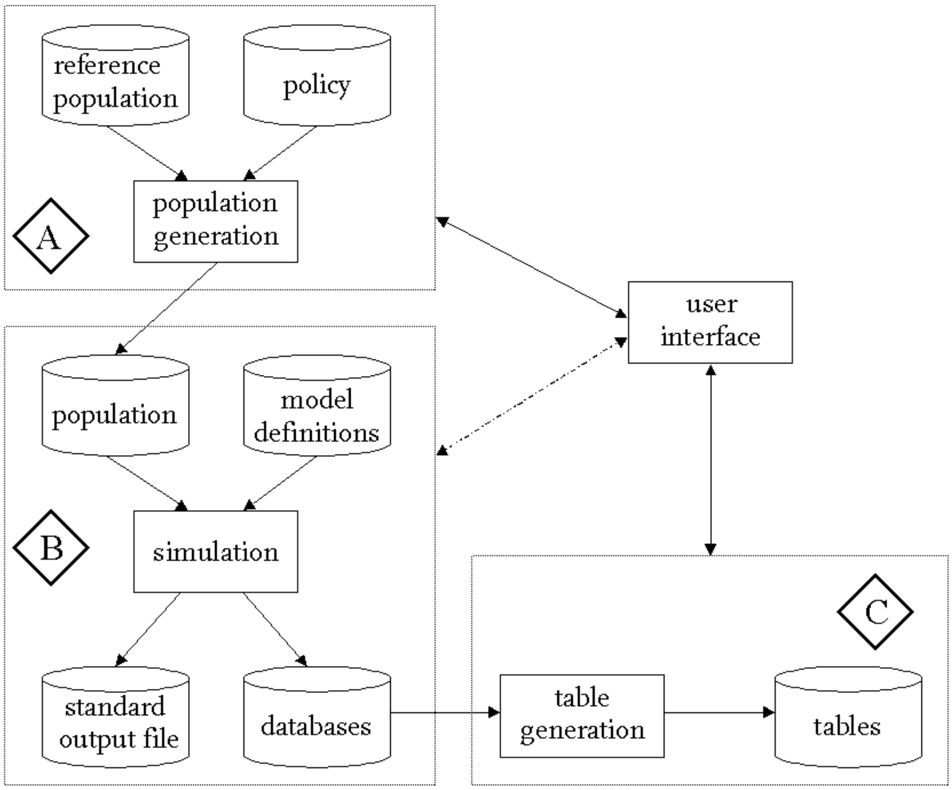


Figure 5.2 The architecture of the MERLIN-system

5.3.1 The User Interface

The User Interface supports the pre- and post-simulation communication with the user. Basically, it provides user-friendly pull-down and “thick” menus that allow the user to specify projects and inspect simulation results. The pre-simulation communication entails the definition of the policies that the user wants to simulate

[18] In this context, a *policy* is defined as an user-defined *scenario for the future* either in terms of changes in the composition and/or characteristics of the population, or in terms of changes in the characteristics of the Tourist-Recreation Product. The generation of the population also comprises the definition of the system and institutional setting within which the simulation units have to make their choices.

and the specification of the simulation settings such as the desired number of simulation runs. The user-defined policies and simulation settings are established in projects. Following the simulation process, the system enters into the post-simulation communication process with the user. Since the standard output files (comprising several indicators of the simulation results at the aggregate level) cannot be inspected within the system, there are no specific tasks for the User Interface with regard to these simulation results. However, based on the generated databases comprising both simulation units and tourist trips, the User Interface does support the generation of tables that show the relationship between two attributes of the generated population and/or the simulated tourist trip patterns.

The second task of the User Interface is to check the entries made by the user to ensure that the proposed changes and project settings are within certain ranges. These (system-defined) ranges prevent the user from specifying policies or project settings that are unrealistic or that would violate the behavioural assumption of the system.

5.3.2 The Simulation Process

The Simulation Process (component B) constitutes the heart of the system because it comprises the main functionality of the system, i.e. (1) simulating the tourist choices of a given population based on a set of decision models; and (2) storing the simulation results in various formats. The interaction between this component and the user is limited: once the user, in interaction with the Population Generator (component A) and the User Interface, has defined a simulation project, the Simulation Process is activated and it completes its task without further notice. There are only two forms of interaction between the Simulation Process and the user. First, during the simulation process, the user is kept abreast of the simulation process via various messages, including the stage of the simulation process in terms of the number of runs that have been completed and the communication of any errors that might occur during the process. Secondly, the user can interrupt the Simulation Process at any time. This way, the simulation process is terminated abruptly and all results are lost.

Given the task of the Simulation Process-component, this component's first field of expertise comprises the sequence of stages of the tourist decision-making process in the *MERLIN*-system (see Figure 5.1). With regard to this task, for each simulation unit, the component collects the relevant condition information from the population database and uses the appropriate decision models to simulate each stage of the tourist decision-making process. During the execution of this task, the

Simulation Process-component stores the simulation results internally, and, if desired by the user, also in databases that include data on the simulation results at the level of the individual as well as the simulated tourist trips. Once this desired number of simulation runs has been executed, the Simulation Process-component generates a standard output file that comprises several indicators of the simulation results at the aggregate level. This file conveys the most important characteristics of the generated population and their tourist trip patterns. In addition, the results of the project under consideration are compared (including tests of significance) to another (previously simulated) project. The second field of expertise of the Simulation Process-component, therefore, is data-base management.

5.3.3 The Population Generator

Based on the user-defined policy settings, the Population Generator (component A) determines the composition and characteristics of the population of simulation units, including the definition of the system and institutional setting within which the simulation units have to make their choices. The Population Generator thus creates a database that comprises the population of simulation units. Chapter 12 (section 12.4) details the available options for specifying policy settings. Also, this chapter discusses how these settings are implemented in the simulation process to represent the scenarios for the future. Suffice it to say, that the Population Generator adjusts a reference population and the reference system and institutional settings to represent the user-defined policy settings.

5.3.4 The Table Generator

Finally, the Table Generator allows the user to inspect the simulation results by generating tables that show the relationship between two attributes of the generated population and/or the simulated tourist trip patterns. These relationships are presented in the form of cross-tables (in the case of two categorical attributes) or tables comparing the mean attribute values for each level of a categorical attribute (in the case of one categorical and one metric attribute). The tables are generated using the two databases that contain the characteristics and tourist choices of each simulation unit in the population, and the characteristics of each simulated tourist trip respectively. This function is only available when the user has requested to generate the required databases and when the simulation process has been completed successfully.

5.4 Conclusion and discussion

This chapter concluded the development of the MERLIN-system by formalising the modelling problem and outlining the architecture of the system. In essence, MERLIN aims to represent and predict the tourist trip pattern for each individual in a particular year, including for each trip, the interrelated choices of the duration, travel party, timing, destination, accommodation, transport mode, date of departure and expenditures, and given the personal and household characteristics of a particular population of simulation units and the system and institutional settings within which these decisions are made. In addition, during the (simulation of the) decision-making process, conditions resulting from previous decisions regarding choice facets of the trip under consideration as well as those on trips with higher profiling priorities are taken into account. By changing the personal and household and system and institutional conditions, scenarios for the future can be evaluated in terms of their impact on annual tourist trip patterns.

The architecture of the MERLIN-system consists of several components that perform specialised tasks. The core component of the system, the *Simulation Process*, simulates the tourist choices that result in annual tourist trip patterns of a given population of simulation units. In addition, this component stores the simulation results in various formats that allow the user to inspect these results. This core component is assisted by several auxiliary components: the *User Interface* supports the pre- and post-simulation communication with the system's user; the *Population Generator* generates the population of simulation units, including the system and institutional settings within which these units operate, that comply with the user defined policy settings; the *Table Generator*, finally, allows the user to inspect the simulation results (stored in two databases) by cross-tabulating the characteristics and tourist choices of the simulation units and/or the simulated tourist trips. Together these components comprise the ability to - in communication with the user - simulate policy options and assess their impact on tourist trip pattern of individuals on an annual basis. In addition, the simulation results are communicated to the user in various ways.

Given this approach to modelling tourist trip patterns, chapters 6 through 11 discuss the empirical building blocks of MERLIN, including a discussion of the data that were collected to derive tourist decision-making rules. The conceptual and empirical building blocks will come together in chapter 12 that will discuss the validation of the system. This chapter also details the available policy options, and evaluates the impact of several of these options on annual tourist trip patterns.

6 Data Collection

This chapter accounts for the data that were collected to calibrate the empirical models. This comprises a discussion of the pilot study that was conducted to produce a list of preconditions for the data collection, the data collection process itself and a discussion of the resulting sample characteristics.

6.1 Introduction

The previous chapters discussed the *MERLIN*-system, including the decomposition of the tourist choice process into a sequence of choice stages. Also, the modelling approaches to describing these stages were identified. Except for the generation of the tourist trip program, i.e. the list of tourist trips that the (potential) tourist will pursue during the year under consideration, tourist choices will be described using decision rules laid down in decision tables. Given these basic principles, data on tourist trip scheduling behaviour, on trip rescheduling behaviour and on the resulting trip patterns is required to induce the rules tourists use to develop their tourist trip patterns. This chapter details the data considerations underlying the *MERLIN*-system. To this purpose, the next section first identifies the data requirements of the system, including an exploration of the possibilities to obtain these data. Next, the results of the pilot study are discussed. This study was conducted to obtain a list of prerequisites for the data collection process. Having identified these preconditions, section four discusses the data collection process in detail. This is followed by a discussion of the resulting sample characteristics and the observed tourist trip patterns. The final section, then, recapitulates the whole process briefly.

6.2 Data requirements of the *MERLIN*-system

The data requirements for developing a comprehensive model of tourist trip patterns like the *MERLIN*-system are as complex as the behaviour under consideration. In short, the *MERLIN*-system requires data on all tourist trips of an individual over a one-year period, comprising both day-trips and overnight holidays, and including information on the most important facets of the travel decision, i.e., the duration, the travel party, the timing, the destination, the accommodation type, the transport mode, the date of departure and the expenditures on each trip. Furthermore, in order to realistically simulate tourist choice behaviour, the system requires information on the initial tourist trip program (i.e., the annual list of tourist trips), the partially profiled and the planned tourist trip patterns, replanning behaviour and the final (observed) tourist trip patterns, including the prevailing conditions at the time of these decisions. Ideally, therefore, one would like an appropriate sample of respondents to keep a diary for a one year period, meticulously recording the various stages and facets of their decision-making processes and the resulting choice behaviour. Unfortunately, given financial, temporal and practical considerations, such an elaborate data collection effort was impossible. Hence, the possibilities of existing data sets in the Netherlands were assessed in terms of their ability to meet the data-requirements of the MERLIN-system. In this process, the Time Budget Survey (Dutch: *Tijdsbestedingsonderzoek* TBO), the National Travel Survey (Dutch: *Onderzoek VerplaatsingsGedrag* OVG), the Day-Trip Survey (Dutch: *Onderzoek Dagrecreatie* OD) and the Continuous Vacation Research (Dutch: *Continu Vakantie Onderzoek* CVO) were considered. It was concluded that combined data on day- and overnight tourist trips of individuals over a one-year period are not available in the Netherlands. Also, information on trip (re-)planning behaviour is not available.

The annual CVO-data sets best approximate our data requirements because it comprises information on the most important personal and household characteristics and on the holidays of a representative sample for (at least) a one-year period. Based on four quarterly measurements, this extensive consumer panel managed by the Dutch Research Institute for Recreation and Tourism (NRIT) and the Dutch Bureau for Tourism (NBT) (Van der Most, 1996), records the annual holiday behaviour, the socio-demographics and holiday-related variables of a panel that is representative of the Dutch population. Furthermore, it provides the possibility to add questions to the quarterly measurements. This way, data regarding day-trips and (re-)planning behaviour of the sample can be obtained.

Based on these constraints and opportunities, it was decided to obtain an entire CVO-data base including information of all overnight trips of approximately 3500 respondents over a one-year period. In addition, estimates of the annual number of day-trips of these respondents could be obtained. Obtaining more detailed data regarding these day-trips was considered impossible given the poor recollection of day-trips after a couple of weeks. As a consequence, the relationship between day-trips and overnight holidays can only be considered at the level of the tourist trip program, i.e. the list of trips that the individual pursues during the year under consideration.

With regard to tourist (re-)planning behaviour, it was decided to confine the extra data collection to the tourist trip program as well. Since this stage of the tourist decision-making process is concerned with generating the annual number of tourist trips, i.e. with time allocation to tourist trips, this process will be referred to as resource re-allocation behaviour.

Before gathering the data required to derive the empirical models, however, the data collection procedures with regard to the tourist trip programs and re-allocation behaviour were tested. Or, in other words, confirmation was needed that respondents would be able to provide reliable data regarding their annual tourist trip programs, and that they would be able to indicate how their annual program would change given a certain change in the decision-making context. The next section reports on the results of a pilot study that was conducted in October 1997 to explore tourist resource (re-)allocation behaviour.

6.3 An experimental exploration of resource (re-)allocation behaviour

In October 1997, a small-scale inquiry into the planning processes for both day- and overnight tourist trips was made. In-depth, semi-structured interviews, including a number of rank-order, allocation and scheduling tasks were conducted to better understand the complex nature of tourist trip planning processes. A small, but diverse sample of respondents served this purpose best. Therefore, a heterogeneous sample, consisting of 22 employees and students of the Eindhoven University of Technology, was selected to participate in a 30- to 40-minute interview that focused on the tourist trip planning process. Potential respondents were contacted by means of the University e-mail network and by verbal invitations to ensure that all social segments of the University population - from the cleaning staff to the professors - were represented. This sampling procedure thus excluded non-working segments of the Dutch population and results in an overrepresentation of higher educated respondents. The advantage, however, is that one might expect higher educated respondents to better phrase their planning processes.

6.3.1 The design of the pilot study

Six different versions of the interview instrument were developed in order to address a large number of topics. The first half of the interview was, however, similar for all respondents. The respondents' answers were tape-recorded (and transcribed verbatim) with their prior permission. This section reports only on those parts of the interview that are relevant to the tourists' resource allocation and re-allocation behaviour with regard to annual tourist trip programs (i.e. the second part of the interview).

At the end of the first part of the interview respondents were asked to recall their holidays and day-trips over the past 12 month to examine whether respondents are able to report their trips over a one year period. Moreover, recalling their past behaviour, respondents were prepared for the (re-)allocation tasks in the second part of the interview. Again, respondents were asked to explicate their choices as much as possible.

The second half of the interviews required respondents to plan their holidays and day-trips for the year to come (1998) under different circumstances and using different response formats. Various tasks were presented to different respondents. The *base-allocation task* examined tourist *time allocation behaviour* and required respondents to draw up their tourist trip program for next year (1998) given 140 free days and assuming their present socio-economic situation. The 140 free days represented all non-working days, including weekends, (public) holidays and the typical Dutch “work-time shortening days” (introduced in the 1980s to redistribute labour more equally among the labour force). Any combination of tourist trips and days at home was allowed, as long as the total number of days would sum up to 140. Thus, the respondents’ preferred level of participation, and their favoured combination of tourist trips could be observed. The response form is presented in Figure 6.1.

Given this initial tourist trip program for 1998, respondents were asked to project their mind to the year 2003: "Some things have changed in our country, and we would like you to reflect on how you would respond to these changes in terms of your holiday and day-tripping behaviour."

Given **140 FREE DAYS** a year (including weekends) for tourist activities, I would like:

Trip type	*Duration*		*Total no. of days*
... day-trips	à 1 day	makes	... days
.. short break(s)	à 3 days	makes	... days
... 1-week holiday(s)	à 7 days	makes	... days
... 2-week holiday(s)	à 14 days	makes	... days
... 3-week holiday(s)	à 21 days	makes	... days
... 4-week holiday(s)	à 28 days	makes	... days
... days at home	à 1 day	makes	... days

		TOTAL	140 days

Figure 6.1 Response form used in the pilot study (original in Dutch)

In order to examine tourist *re-allocation behaviour*, including the task and information load of various changes in the tourists' decision-making context four different *scenarios* were developed, each addressing two dimensions of the decision-making context. The first dimension covered the influence of working hours on annual tourist trip programs (*demand side*). Shorter working weeks were introduced in the 1980s to reduce unemployment in the Netherlands. In their battle against unemployment and in favour of the protection of physical and mental well-being, the associations of labour unions increased their demands and started to negotiate 36-hour and even 4-day working weeks halfway through the 1990s (Raaijmakers, 1997). Upon success, employees were either granted extra *ATV*-days, or the working week was reduced to four (or four and a half) days. Recently, however, the labour shortage in certain industries and certain parts of the public sector has forced employers to aim for the reintroduction of the 40-hour working week. Given these contradictory developments it was decided to examine the effect of both an increase and a decrease of *ATV*-days on annual tourist trip programs. Respondents were told that, due to new collective labour agreements, the number of *ATV*-days had been decreased or increased by 12 days. Given the 140 free-days in the base-allocation task, the decreased free days-scenario came down to 128 days free day and more income. In contrast, the increased free days-scenario came down to 152 days free days and less income; all other things the same.

The second dimension of the scenarios covered changes in the Tourist Recreation Product (TRP; *supply side*). Inspired on the discussion note "Leisure Outline 2020" (Dutch: *Recreatieschets 2020*) by the Ministry of Agriculture, Nature Conservation and Fisheries, two development directions for the TRP were advanced (Ministry of LNV, 1997). The "work"-scenario was defined as (all other things the same):

- an increased number of recreation areas with improved accessibility;
- an increased number of bungalow accommodations like "Centre- Parcs";
- improved public transport services; and
- cars and aeroplanes more expensive.

The "consumption"-scenario was presented as (all other things the same):

- green areas shredded and privatised;
- increased road traffic and dispersed living areas;
- more private leisure clubs; and
- an emphasis on culture in cities.

Combining the demand and supply side changes produces four scenarios. Each respondent was presented with one of these scenario’s and asked to indicate the numbers of day-trips and holidays (s)he would make given the initial plans for 1998. Again, respondents were asked to explicate their ideas and choices as much as possible.

6.3.2 The result of the pilot study

The interview protocols reveal that respondents usually do not find it difficult to indicate their (preferred) holidays for next year, probably due to rather stable trip patterns. With regard to the planning of day-trips, however, respondents find it very difficult to assess the number of day-trips they (want to) make. During the interviews it became clear that if the interviewer gave a lot of examples (shopping for fun, cycling, walking, picnicking, visiting a town, theme parks, swimming, eating out for a couple of hours, etc.), respondents considered it easier to indicate for each trip type how many times they would (want to) do this per week, per month or per year. A rough estimate of the total number of day-trips could then be acquired by counting the estimates for each trip type the respondent would pursue. Finally, respondents facing major changes (graduation, first/new job, marriage/divorce, new-born children or retirement) were less confident about next year's trip program.

With regard to the scenario or re-allocation task, respondents sometimes reasoned out of their own personal situation in five years. As an example, several older respondents took into account the fact that in five years time their heath conditions might have changed. With regard to the two dimensions in the decision-making context, respondents had a hard time trying to imagine the changes to the TRP and the consequences for their own behaviour. Also, several respondents indicated that they would plan their holidays without taking into account the actual condition of the TRP in detail. The majority of changes in the trip program could be attributed to the in- or decrease in the number of free days, the changes in the personal situation in five years time and sometimes to the changes in the transport system.

6.3.3 Preconditions for data collection on tourist trip programs

Given the results of the interviews, six preconditions for the data collection approach were defined. First, given the high information load of both the basic allocation task and the scenario task, it was concluded that "experienced" respondents should be consulted in any future data collection procedure. In this context, "experienced" refers to the respondents' knowledge of the definitions of tourist trips, their familiarity with their own tourist behaviour, and their skills with regard to filling out questionnaires.

Second, in addition to consulting experienced respondents, the questionnaire should provide warming up tasks to assist the respondent in recalling his or her tourist trips during a specific period. This applies in particular to day-trips that can easily escape a respondent's memory due to the short-lived nature of these trips.

Third, with regard to the scenarios it was concluded that they should not be positioned in the future but rather be defined as changes to the present situation. For example, given the present (or last year's) trip program, respondents should be asked to indicate what they would have done differently given certain changes. This will prevent respondents from pondering upon their lives in the future, and produce more workable data.

The fourth precondition for the data collection approach is the need to present scenarios that correspond to the respondents' personal situation. For example, retired people should not be presented with a scenario granting them more free days. Also, changes in the decision-making context should not be too dramatic because this may confuse respondents.

With regard to the scenario dimensions, it was concluded that the adopted approach lends itself better to gathering information on possible responses to changes on the demand side of the tourist industry rather than changes on the supply side because respondents find it difficult to imagine these changes or they indicate that this does not affect their planning process. Consequently, the fifth precondition would be to limit the presented changes to the demand side as much as possible.

Finally, given the time and efforts required to contemplate the scenarios, respondents should not be asked to consider more than one scenario. If several scenarios are to be presented to a particular segment of respondents, they should be distributed randomly over the respondents of that segment.

6.4 Data collection: Continuous Vacation Research (CVO)

Based on the results in the pilot study, the allocation task was adjusted and simplified and the best scenarios were selected and elaborated. Also, a suitable data-collection procedure was required to gather reliable data on tourist trip programs including information on both day-trips and holidays over a one-year period.

In accordance with the first precondition that resulted from the pilot-study, the experienced "CVO-panel" was used. Data on the holidays of 3562 respondents

(0-99 years) were obtained by acquiring the 1998 CVO-data (a "CVO-year" runs from December through November of the following year). As mentioned before, this set also includes the most important socio-demographic variables of the respondents. In addition, data was needed on the respondents' day-tripping behaviour and their behavioural responses to changes in the decision-making context (at the level of tourist trip programs). This data was acquired by participating in the fourth quarterly CVO-measurement in December 1998. The following questions were presented to *adults only (16+ years)*.

In accordance with the need to include "warm-up tasks" (precondition 2), respondents were first asked to indicate the number of day-trips they had made during the past year. Broken down into 10 trip types, possible day-trips included swimming/sun-bathing, cycling/walking /skating, sporting trips, visiting or supervising sporting events, go for a drive, visiting a theme park/Zoo, visit a city/town/museum, visit an event (fair, show, exhibition, etc.), fun-shopping and going out (movie, restaurant, café, and so on). A day-trip was defined as an "outdoor trip from the residential location for recreational purposes, involving at least four successive hours without spending a night". Apart from "never", the response form allowed respondents to indicate the number of trips for each trip type per week, per month or per year.

∏ From December 1997 through to November 1998 inclusive I did not make any day-trips nor did I make any short or longer holidays.

∏ From December 1997 through to November 1998 inclusive I did make day-trips and/or short or longer holidays, namely:

.... Day-trips (min. 4 hours)	
.... Short breaks (2-4 days);	of which ..... in private tourist accommodation
.... Medium long holidays (5-8 days)	of which ..... in private tourist accommodation
.... Extended holidays (9-15 days)	of which ..... in private tourist accommodation
.... Long holidays (16-28 days)	of which ..... in private tourist accommodation
.... Extra long holidays (29+ days)	of which ..... in private tourist accommodation

Figure 6.2 Response form used in the data collection process in December 1998 (original in Dutch)

Next, respondents were asked to indicate the number of day-trips and holidays they had made during the past year (see Figure 6.2; the question on the number of trips in private tourist accommodation was added to stimulate respondents to also record these trips as part of their "normal" tourist trip pattern). With regard to the number of day-trips, respondents were encouraged to use the previous question on day-trips as a mnemonic devise. Respondents were provided with a response form that allowed them to indicate the number of day-trips (of a minimum 4 hours), short breaks (2-4 days), medium long holidays (5-8 days), extended holidays (9-15 days), long holidays (16-28 days) and the number of extra long holidays (29+ days)[19]; stepwise instructions were included to guide respondents through this task.

Given their own list of tourist trips in 1998 (trip program), respondents were asked to answer the question "What would you have done in 1998 if.....*scenario*"; this question satisfied precondition 3. The scenarios, or changes in the decision-making context varied (1) between different socio-economic groups (or segments) within the panel to avoid the problem of unrealistic changes in the decision-making context (precondition 4); and (2) within homogeneous socio-economic groups to allow for multiple scenarios to be included without overburdening individual respondents (precondition 6). Group-specific scenarios were randomly assigned based upon the respondents' months of birth. Respondents were presented with the same response form (Figure 6.2) to indicate their trip program given the specific change in the decision-making context.

Five socio-economic groups were distinguished based on two important constraints in the tourist decision-making process: (1) the availability of *time*, as indicated by the respondents' (or the respondents' households') working and educational situation; and (2) the availability of *financial resources* (see Figure 6.3). Group 1 included respondents who had been working at least 30 hours per week in 1998. Correspondingly, group 2 was comprised of respondents who did not work full-time themselves, but who were restrained by their partners' commitment to a (minimum) 30-hour working week. The respondents in group 1 and 2 are assumed to have significant time constraints, but no, or few, financial limitations. The scenarios for these groups systematically vary the most likely changes in working hours for the near future (demand side; precondition 5).

Group 3 included the students (and pupils aged 16 and over) in the CVO-panel. These respondents are assumed to have rather flexible time schedules, and few financial resources. The scenarios presented to this group concern the so-called "Student Public Transport Ticket" (Dutch: *Studenten OV-jaarkaart*). On January first, 1991, the Ministry of Sciences, Education and Culture, introduced a ticket that allowed students with a government educational grant (Dutch: *Studiefinanciering*) to use all forms of public transport 365 days a year. The educational grant was decreased with approximately NLG 60 per month to pay for this ticket. Compelled by government cuts throughout the years, the conditions

[19] In contrast to the pilot study, the duration of the holidays was expressed in ranges of days rather than in a "number of weeks".

of this ticket have been changed several times. In 1998, students living at home were allowed to use the ticket during working days (Monday morning through to Friday 19.00 hours), excluding Christmas, Easter, Queensday, Liberation Day, Ascension Day, Whit Monday and the months of July and August.

Group 1: People working at least 30 hours per week (in 1998)
Scenario A [Jan, June or Nov]: Due to new collective labour agreements you had had 12 extra free days ; your salary had been the same;
Scenario B [Feb, July or Dec]: Due to new collective labour agreements you had had 12 free days less; your salary, however, had increased by 5%;
Scenario C [March or Aug]: Due to new collective labour agreements you had worked 4 days a week; Wednesday had been your fixed free day; your salary had been the same;
Scenario D [April or Sept]: Due to new collective labour agreements you had worked 4 days a week; Monday had been your fixed free day; your salary had been the same;
Scenario E [May or Oct]: Due to new collective labour agreements you had worked 4 days a week; Friday had been your fixed free day; your salary had been the same.

Group 2: People working less than 30 h.p.w. with a partner working at least 30 h.p.w.
Scenario F [Jan to June]: Due to new collective labour agreements your partner had had 12 extra free days; his/her salary had been the same;
Scenario G [July to Dec]: Due to new collective labour agreements your partner had had 12 free days less; his/her salary, however, had been increased by 5 %.

Group 3: Students and pupils
Scenario H [Jan, March, May, July, Sept or Nov]: In 1998 you had been allowed to use your Student Public Transport Ticket only for trips to your educational institution; in addition, you had had 7 days on which you could use this ticket for private trips; these additional days could not take place on Mondays and Fridays;
Scenario I [Feb, April, June, Aug, Oct or Dec]: In 1998 you had been allowed to use your Student Public Transport Ticket as it was introduced in 1991, viz. free use of all public transport modes within the Netherlands on all days.

Group 4: Pensioners, early retirement (Dutch: *VUT*) and people of independent means
Scenario J [Jan, March, June, July, Sept or Dec]: Due to changes in the Dutch tax and social legislation, your net household income had been 15% higher in 1998;
Scenario K [Feb, April, May, Aug, Oct or Nov]: In 1998 you had had 7 (extra) days on which you could travel by train free of charge (in case you already had 7 free travel days on your “benefit hours ticket” (Dutch: *Voordeelurenkaart*) you would have had 14 free travel days); every two months (or every month) you could have travelled using one of your free travel days, and the remaining day(s) was (were) to be chosen freely; these free travel days could not take place on Mondays and Fridays.

Group 5: Respondents who do not belong to any of the above groups (“other”)
Scenario L [Jan, March, May, July, Sept or Nov]: In 1998 you had had 7 days on which you could travel by train free of charge; every two months you could have travelled using one of your free travel days, and the remaining day was to be allocated freely; these free travel days could not take place on Mondays and Fridays
Scenario M [Feb, April, June, Aug, Oct or Dec]: Due to changes in the Dutch tax and social legislation, your net household income in 1998 had been 15% higher.

Figure 6.3 The 13 scenarios presented to the respondents [month of birth]

In contrast, students living in student rooms (in or near the city of their educational institution) were allowed to use the ticket from Friday 19.00 hours through to Sunday evening, and on the holidays previously mentioned. The cut on the education grant due to the Public Transport Ticket amounted to approximately NLG 100 per month. The scenarios presented to the students either improved or diminished the conditions of the Student Public Transport Ticket for 1998.

Group 4 included panel members who are either retired or living off their investments. Group 5, finally, was comprised of people that did not belong to any of the other groups. Respondents in this group and their possible partners did not work full-time (less than 30 hours), they were unable to work and/or unemployed. These groups were assumed to have no or few time constraints. Financial consideration, old age and delicate health, on the other hand, may dominate the tourist decision-making processes of these people. The scenarios presented to these respondents relief some of the financial constraints (demand side). In addition the scenarios encourage the use of public transport[20] (supply side).

6.5 Sample characteristics

The 1998 CVO-data set and the information obtained from the additional resource (re-)allocation questions provide the empirical basis from which the *MERLIN*-system is built. The purpose of this final section is to describe this sample of 3562 respondents in terms of the socio-demographic characteristics as well as the tourist trip patterns.

Since the additional questions were presented to adults (16+ years) only, the discussion will contrast the characteristics of three segments of the sample: (1) the entire 1998 CVO-sample, including both children and adults (3562 respondents); (2) the adult segment of the 1998 CVO-sample consisting of the 2836 respondents aged 16 years and over; and (3) the adult respondents whose responses to the additional questions were usable for calibrating the resource (re-)allocation model (1530 respondents; the other respondents were not selected due to missing values and invalid responses). The emphasis of this section will be on comparing the adult-segment and the “usable allocation responses” in order to establish the representativeness of the latter segment on which the utility-based model for time allocation to tourist trips will be calibrated.

[20] In the Netherlands, people aged 65 years holding a “benefit hours ticket for trains” (Dutch: *Voordeelurenkaart*) can use the train free of charge on 7 days within the period of one year. Inspired on this system, the seven free travel days were also introduced to the respondents in group 5.

Table 6.1 Socio-demographic characteristics of the 1998 CVO-panel

	1998 CVO-panel	1998 CVO-panel ≥16 years	usable allocation responses
$N_{respondents}$	3562	2836	1530
Household income			
< NLG 35,000	1112 (31.2%)	948 (33.4%)	437 (28.6%)
NLG 35-55,000	1549 (43.5%)	1180 (41.6%)	661 (43.2%)
> NLG 55,000	901 (25.3%)	708 (25.0%)	432 (28.2%)
Age (sd)	37.9 (21.4)	45.7 (16.5)	43.2 (15.7)
Education level			
- low	1909 (53.6%)	1184 (41.7%)	551 (36.0%)
- intermediate	1030 (28.9%)	1030 (36.3%)	575 (37.6%)
- high	623 (17.5%)	622 (21.9%)	404 (26.4%)
Household size (sd)	3.01 (1.40)	2.70 (1.31)	2.77 (1.32)
Holiday region [a]			
- North	1262 (35.4 %)	1001 (35.3%)	517 (33.8%)
- Mid	1016 (28.5 %)	805 (28.4%)	455 (29.7%)
- South	1284 (36.0 %)	1030 (36.3 %)	559 (36.5%)
N° females	1875 (52.6%)	1528 (53.9%)	800 (52.3%)
Car(s) in hh	2922 (82.0%)	2307 (81.3%)	1275 (83.3%)
Ski	338 (9.49%)	280 (9.87%)	169 (11.0%)
Permanent tourist acc.	255 (7.16%)	215 (7.58%)	97 (6.34%)
Non-permament tourist accommodation	1279 (35.9%)	974 (34.3%)	565 (36.9%)
Living in urban area	1368 (38.4%)	1147 (40.4%)	615 (40.2%)
Work (paid job)	1580 (44.4%)	1580 (55.7%)	946 (61.8%)
Working hours p.w. (sd)	14.48 (17.88)	18.18 (18.28)	20.21 (18.13)
Child in household			
- no	1875 (52.6%)	1875 (66.1%)	964 (63.0%)
- yes, 6-17 years	1273 (35.7%)	758 (26.7%)	430 (28.1%)
- yes, only 0-5 years	414 (11.6%)	203 (7.16%)	136 (8.89%)

[a] In the Netherlands, the summer, autumn and spring holidays of primary and secondary schools are spread regionally to avoid congestion problems and increase occupancy rates.

Table 6.1 presents the socio-demographic characteristics of the (segments of the) 1998 CVO-panel. As expected, in contrast to the entire CVO-panel, the adult-segment and the "usable allocation responses" are older, they have smaller household sizes due to the absence of children, they have higher levels of education and participation in working activities and they live in urban areas more frequently. These differences can all be reduced to the age-selection criterion.

Comparing the adult-segment to the “usable allocation responses”, it is evident that the higher income and education brackets are over-represented in the latter segment. Also, the percentage of working people and the average number of working hours per week are higher in the “usable responses”-segment. With regard to the other socio-demographic characteristics, the “usable allocation responses”-segment is somewhat younger than the entire adult segment, and owners of tourist commodities (skis and non-permanent accommodations) and households with young children are slightly overrepresented in comparison to the “adult” segment. Differences are, however, small, and should not cause any serious problems.

Table 6.2 provides an overview of the tourist trip patterns of the (three segments of the) 1998 CVO-panel. Comparing the entire 1998 CVO-panel to its adult segment reveals that, on average, (households with) children have higher participation rates for holidays, but that, on the other hand, they have lower trip frequencies and annual expenditure rates. With regard to annual expenditures on holidays, it is evident that these differ greatly across the population because for all three segments the standard deviation exceeds the mean value. Comparing the annual expenditures on holidays of the adult segment to the “usable allocation responses”, the latter segment on average allocates more financial resources to (overnight) tourist trips than the first. This is consonant with the observation that the higher income and educational brackets are slightly overrepresented in the “usable allocation responses”-segment.

Comparing the holidays reported by the two adult segments reveals that the participation rates of the “usable allocation responses” are similar to those of the entire adult-segment. This also applies to the number of medium long, extended and long holidays (per person per year), and the overall number of holidays per year. However, the average number of short breaks is (seriously) overestimated, whereas the average number of extra long holidays is (seriously) underestimated. Both sample and response biases could account for the latter observations. Validation of the obtained data for the “usable responses”-segment at the individual level is reported in Chapter 7.

The day-trip frequencies of the “usable responses”-segment cannot be compared to those of the other segments. Comparing the data to national statistics, however, reveals that the reported participation rates and trip frequencies are probably underestimated. According to the most recent Day-Trip Survey (CBS, 1997), people on average make 25.9 day-trips per year (based on a 4-hour definition), compared to 14.9 day-trips for the “usable allocation responses”.

Chances are that the respondents have seriously underestimated their day-trip frequencies. However, since reference data regarding day-trips are not available for the 1998 CVO-panel members, *MERLIN* will not be adjusted for these possible data imperfections.

Table 6.2 Tourist trips of the 1998 CVO-panel

	1998 CVO-panel	**1998 CVO-panel ≥16 years**	**usable allocation responses** [a]
$N_{respondents}$	3562	2836	1530
- making holidays	2851 (80.0%)	2239 (78.9%)	1202 (78.6%)
- making day-trips	-	-	1076 (70.3%)
$N_{holidays}$	7121	5808	3242
- short breaks (p.p.)	2522 (.7080)	2063 (.7274)	1396 (.9124)
- medium long holidays (p.p.)	2027 (.5691)	1661 (.5857)	826 (.5418)
- extended holidays (p.p.)	1438 (.4037)	1171 (.4129)	606 (.3961)
- long holidays (p.p.)	961 (.2698)	757 (.2669)	363 (.2373)
- extra long holidays (p.p.)	173 (.04857)	156 (.0550)	48 (.0314)
Average N° holidays	2.00	2.05	2.12
- per participant	2.50	2.59	2.70
Duration holidays (sd)	9.13 days (7.57)	9.12 days (7.71)	-
Expenditures on all holidays (sd)	NLG 1425.52 (1871.94)	NLG 1593.84 (2027.86)	NLG 1668.23[b] (2166.23)
N° of day-trips	-	-	22865
- average N° day-trips	-	-	14.9
- per participant			21.3

[a] These data are based on the responses to the additional (re-)allocation questions

[b] This average is based on the holidays reported by these respondents in the regular 1998 CVO-data bases, and not on the answers to the additional questions

Based on these observations it can be concluded that the socio-demographic composition of the "usable allocation responses"-segment does not cause many problems, but that the trip patterns of these respondents are most likely to underestimate the number of day-trips and extra long holidays and to overestimate the number of short breaks. Chapter 12 will discuss how the *MERLIN*-system will weight the predictions based on these data to counterbalance the data deviations for *holidays*.

6.6 Conclusion

This chapter rendered account of the data that are used to build the *MERLIN*-system. First, given the methodological and conceptual considerations presented in the previous chapters, the data requirements of the system were defined. Secondly, the Continuous Vacation Survey (*CVO-panel*), was selected to provide information on the holiday behaviour and the socio-demographics of a sample of approximately 3500 people, representative of the Dutch population. It was concluded that this survey best approximated the data requirements of the *MERLIN*-system. In addition, it was possible to present supplementary questions to the members of this panel. This way, the CVO-data on observed holidays were complemented with information on day-trips and re-allocation behaviour.

First, however, a pilot study was conducted. Based on this study, six preconditions for the data collection were identified. Given these preconditions, this chapter outlined the approach to obtaining the data required for the *MERLIN*-system. Finally, the obtained data were described in terms of the socio-demographic characteristics of the 3562 respondents and observed tourist trip patterns. In particular, the discussion focussed on comparing the adult (16+)-segment (comprising 2836 respondents) and the 1530 usable responses to additional questions that were returned by this sample. It was concluded that in comparison to the whole adult-sample, the socio-demographic composition of "usable responses"-segment did not deviate significantly. However, the observed tourist trips of these usable responses proved to underestimate the number of day-trips and extra long holidays and to overestimate the annual number of short breaks.

The subsequent chapters elaborate the empirical models that are derived from the data described in this chapter. The preparation of the data is discussed in each chapter separately.

7 Tourist Time Allocation Choices

This chapter describes time allocation to tourist trips. This process is interpreted as a two-staged problem. First, the individual decides whether or not (s)he would like to participate in day- and/or overnight tourist trips. Next, conditional upon this choice, the individual may allocate a certain amount of time to the various tourist trips. Given the available data, two (sets of) models are presented, one for young people and one for people aged 16 years and over. The latter model also captures tourist re-allocation behaviour induced by changes in the decision-making context.

7.1 Introduction

This chapter is the first in a series that, based on the conceptual considerations and data that were discussed in the previous chapter, presents the empirical input into the *MERLIN*-system. This first empirical chapter describes the selection of the set of tourist trips an individual will pursue during a year (i.e. the annual tourist trip program including both day- and overnight tourist trips). Based on the conceptual model presented in chapter 2, the selection of the annual set of tourist trips is interpreted as a time allocation problem in which the decision to pursue tourist trips is traded-off against other time consuming activities. In addition, tourist resource allocation is conceptualised as a two-stage process in which, first, the (potential) tourist decides whether or not to participate in day- and/or overnight tourist trips - a discrete choice. And second, given the decision to participate in day- and/or overnight tourist trips, the individual allocates a certain amount of time to tourist trips - a more continuous allocation process. Chapter 3 reviewed several utility- and non-utility-based modelling approaches and concluded that a hierarchical logit model would describe the assumed two-staged nature of the tourist resource allocation process best. Thus, in order to develop a model that can generate annual tourist trip programs, this chapter models tourist time allocation processes using a hierarchical logit approach.

Given the adopted data-collection approach (see previous chapter), data are available on the number of day- and overnight tourist trips for those aged 16 years and over. Also, some information regarding re-allocation behaviour as a result of relieving or reinforcing constraints in the tourist's decision-making context is available for this sample. These data are not available for those aged 0-15 years because this part of the sample did not answer the questions on day-trips and on resource re-allocation; in the 1998 CVO-data set, information is available only on the holidays of this group. As a consequence, *MERLIN* only considers the overnight trips of individuals between 0 and 15 years and different models are developed to describe tourist time allocation processes for children and adults.

This chapter is organised as follows. First, the following section outlines a general approach to modelling tourist time allocation processes to day- and overnight trips, including re-allocation behaviour at the level of annual trip programs. This discussion also addresses the preparation of the data. Next, the model of tourist time allocation for those aged 16 years and over is introduced ("adult"-model). Finally, the model for younger people is presented ("child"-model). As mentioned before, due to the lack of appropriate data, this model does not include time allocation to day-trips nor does it cover re-allocation behaviour. In conclusion, this first chapter on empirical results summarises the major findings on tourist time allocation.

7.2 General modelling approach and data considerations

The subsequent subsection first elaborates the participation choice by exploring the relationship between day-trips and holidays. Next, a hierarchical logit approach is developed that represents the two-staged tourist time allocation process. Finally, the general modelling approach to tourist time allocation processes is discussed and the data preparation is accounted for.

7.2.1 The relationship between day-trips and overnight holidays

Little is known about the interdependencies, if any, between day-trips and holidays in the decision-making process of the tourist. Theoretically, day- and overnight trips should be related because any consumption decision necessarily involves some allocation of time and money - resources that are ultimately finite (Holbrook & Lehmann, 1981). Given the absence of models that focus on the allocation of time across day- and overnight trips, the pilot study (see previous chapter) also explored the relationship between these trips (reported in: Van Middelkoop *et al*., 1999). According to these analyses, day-trips and holidays of 5 days or more are competitors (as indicated by a negative correlation between the time spent on these trips during a particular year), while short breaks (2-4 days) and longer holidays are complements (positive correlation); short breaks and day-trips, finally, proved to be independent (no significant correlation between the time allocated to these trips).

Based on these analysis there is no need to distinguish between short breaks and longer holidays in the participation choice in order to do justice to the competitive relationship between day-trips and longer holidays because the time allocated to short breaks increases as the time allocated to longer holidays increases. The *MERCIN*-system therefore distinguishes four participation choice alternatives: on an annual basis a (potential) tourist may decide to pursue (1) day-trips only; (2) holidays (2+ days) only; (3) both day-trips and holidays; or (4) no tourist trips at all. This classification allows one to examine any substitution effect between different kinds of tourist trips.

7.2.2 A general modelling approach to tourist time allocation processes

In contrast to the participation choice, *allocation of time to tourist trips* is viewed as a more continuous allocation process. In this process the tourist decides on how to distribute the available amount of time across the available alternatives. Given the data presented in the previous chapter, these alternatives include day-trips (of a minimum 4 hours), short breaks (2-4 days), medium long holidays (5-8 days), extended holidays (9-15 days), long holidays (16-28 days) and extra long holidays (29+ days). In case of the annual set of tourist trips, the tourist has 365 (or 366) days at his or her disposal. If "one day" is selected as the unit of analysis, this choice can also be viewed as a process of allocating 365 units to either one of the tourist trips or to a non-tourist activity (such as working, going to school or visiting friends and relatives). The resulting general model for tourist time allocation processes for day- and overnight trips is presented in Figure 7.1.

Using a utility-maximising approach, it is assumed that the utility derived from allocating a certain amount of time to a particular (tourist or non-tourist) activity is composed of a structural part and an error component. When the error components are assumed to be Gumbel distributed, logit models can be used to model the time allocation process. Given the two-staged tourist time allocation process, a *nested logit model* (see chapter 3) would represent the hierarchical structure of the choice process best. In this case, the structural utility of an *trip alternative* is assumed to be composed of (1) a base utility represented by a constant; (2) socio-economic characteristics of the individual (representing personal and household constraints and opportunities); and (3) the condition of the decision making context (scenarios; context effects capturing possible resource re-allocation behaviour). The *participation choice* is also assumed to be affected by the socio-economic characteristics of the (potential) tourist and the scenarios. In addition, the model for the participation choice includes an extra parameter that expresses the ratio of the scales between choices at the participation and the trip choice levels.

Since the exogenous variables in the participation and the trip models do not vary across the choice alternatives, their effects can only be established by using polytomous logit models in which all independent variables are alternative specific and in which one alternative is used as the reference. In practice, the utility of the reference alternative is fixed at zero, and all other effects are estimated relative to this base.

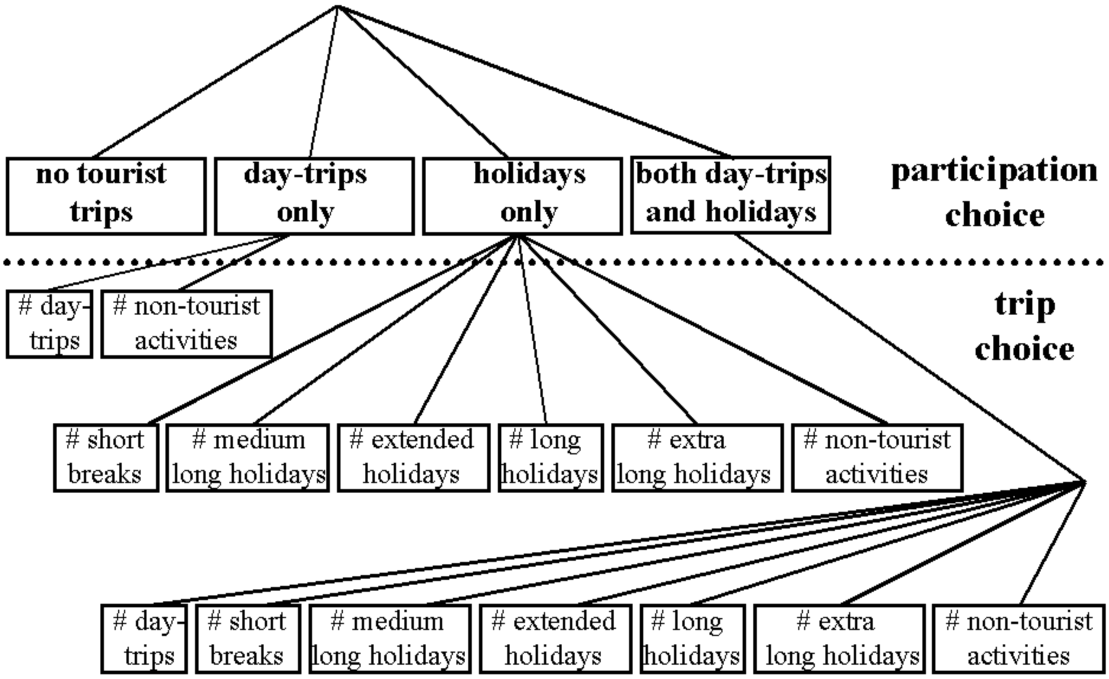


Figure 7.1 A general model for time allocation processes for day- and overnight tourist trips (# : the time (expressed in the number of days) allocated to a particular trip type)

7.2.3 Data considerations

Time allocation data are available for six tourist trip types (see previous chapter). Also, re-allocation behaviour under several scenarios (Figure 6.3) was recorded. In total, 1530 useful questionnaires were returned by the adult members of the 1998 CVO-panel (53.9% of the sample).

To approximate the time that people allocate to the various trips, the trip frequencies are converted to "the number of days allocated to a tourist trip". By definition, day-trips take one day; other trips are set at 3 days (short breaks), 7 days (medium long holidays), 12 days (extended holidays), 22 days (long holidays) and 30 days for extra long holidays. The time allocated to non-tourist activities is calculated by deducting the numbers of days spent on tourist trips from 365 days.

The questions that were added to the CVO-measurement (see chapter 6) required respondents to recall their tourist choices over a one-year period in 1998. In order to validate the obtained data, the responses to the additional questions are collated with the information on the *holiday*-behaviour of the same sample available from the standard CVO-data set[21]. Compared to the 1998 CVO-data set, 68.9% of the respondents have reported the exact number of annual holidays in the additional question; 90.5% of the respondents have reported the correct number of holidays plus or minus one; and 95.4% have done so plus or minus two holidays. Regressing the number of holidays in the CVO-data set (independent) onto the number reported in the additional questions (dependent) further subscribes to the validity of the responses to the additional questions: the intercept only differed slightly but significantly from zero (0.117; t = 2.4) and the slope of the number of holidays in the 1998 CVO-data set nearly equalled one (0.994; t-value with respect to (the deviation from) unity = -.3634; R = .839; (adjusted) R-Square = .703). Finally, the number of days allocated to overnight holidays according to the 1998 CVO-data set (independent) was regressed onto the number of days allocated to tourist trips based on the additional questions (using the above mentioned conversion; dependent). The results of this regression analysis indicate that the intercept is slightly higher than desired (intercept = 1.713; t = 5.94), and that the slope is slightly smaller than unity (slope = .854; t-value with respect to (the deviation from) unity = -12.84; R = .887; (adjusted) R-Square = .787).

Although the above analysis indicate the responses to the additional questions to be rather reasonable in terms of the annual number of holidays and days allocated to these trips, this does not mean that this also applies to the individual tourist trips. A simple comparison of the average number of short breaks for the 2836 adults in the 1998 CVO-panel (.7274) and that for the 1530 usable (re-) allocation forms (.9124), for instance, reveals that the data that are used for the estimation of the time allocation models overestimate the number of short breaks (see Table 6.2). On the other hand, the numbers of extra long holidays are (severely) underestimated. Chapter 12 discusses how *MERCIN* will weight the predictions based on the models that are developed in this chapter to counterbalance these data irregularities. First, however, the next two sections present the "adult"- and the "child"-models.

7.3 The "adult"-model for tourist time (re-)allocation

This section presents the model of tourist time allocation to day- and overnight trips for adults (> 15 years). This model is used in the *MERCIN*-system to generate the annual set of tourist trips, i.e. the tourist's trip program, for each

[21] The CVO holiday data include the annual number of holidays per respondent, as well as the exact duration of each holiday in days. Given the nature of this panel, data on the number of *day-trips* is not available.

individual given his or her personal and household constraints and opportunities. The model also includes parameters that indicate how people would re-allocate their resources given certain changes in the decision-making context.

7.3.1 Overall time allocation to tourist trips by adults

First, differences in the total amount of time allocated to tourist trips between the participation groups were examined. Table 7.1 displays the average number of days allocated to all tourist trips for each of the four participation groups in 1998 (excluding the scenarios).

Table 7.1 The average number of days allocated to tourist trips in 1998 per participation group (after conversion)

	No tourist trips	Day-trips only	Holidays only	Both day-trips and holidays	Total
1998 (N)	0.00 (195)	20.23 (133)	20.96 (259)	43.94 (943)	32.39 (1530)

The averages differ significantly between the four participation groups (F=183.5; p=.000). The averages for the “day-trips only” and the “holidays only” groups, however, do not differ significantly (F=0.127; p=.721). The “both day-trips and holidays” group allocates more than twice as much time to tourist trips than the “day-trips only” and “holidays only” groups. This suggests that tourists who allocate little time to tourist trips are more likely to select either day- or overnight tourist trips.

7.3.2 The tourist time allocation models for adults

Ideally, a joint participation-trip choice model should be estimated from which the Inclusive Values (IVs) of the three participation nests are derived. Unfortunately, however, simultaneous estimation was not feasible because of the large number of alternative specific parameters and the capacity limits of the available software. Alternatively, therefore, three separate Multinomial Logit (MNL-) models were estimated for the “day-trips only”, the “holidays only”, and the “both day-trips and holiday” nests. The time allocated to non-tourist activities was selected as the base alternative in each nest. The socio-economic variables that were entered into the models are listed in Table 7.2.

Table 7.2 The socio-demographic variables that are used to explain participation and trip choices [cont. = continuous]

Variable	Levels	Coding	N/average
Net household income	less than NLG 35,000	[1,0]	437
	NLG 35,000-55,000	[0,1]	661
	more than NLG 55,000	[0,0]	432
Educational level	low	[1,0]	551
	intermediate	[0,1]	575
	high	[0,0]	404
Age	16-30 years	[1,0]	343
	31-55 years	[0,1]	864
	56+ years	[0,0]	323
Gender	female	[1]	800
	male	[0]	730
No. of people in household		[cont.]	2.77
Children in the household	yes	[1]	566
	no	[0]	964
Residence	urban	[1]	615
	rural	[0]	915
(a) Car(s) in household	yes	[1]	1275
	no	[0]	225
Accommodation with a permanent place	yes	[1]	97
	no	[0]	1433
Accommodation without a permanent place	yes	[1]	575
	no	[0]	955
Work	employed	[1]	946
	no job	[0]	584
Working hours/week		[cont.]	20.21

Seven of the 13 scenarios (see Figure 6.3) were entered as alternative specific dummy variables[22]. Parameters that were not significant at the 95% confidence level were excluded. Next, for each respondent individually, the IVs for each nest were calculated. The IV of a nest represents the maximum utility of the alternatives in the lower level nest. In the higher-level participation choice model, the parameter of the IV expresses the ratio of the scales between the participation and the (time allocated to the) trip choices. Finally, the participation choice model was

[22] Based on exploratory analysis, the effects of 6 of the 13 scenarios were not significant or did not seem plausible (i.e., the free-days scenarios C and D, the transport scenarios H, K and L and the income scenario M). Perhaps it was too difficult to imagine these changes and/or the number of observations was too low. These scenarios are not included in the following analysis.

estimated using the "no-tourist trips" alternative as a reference. This model included alternative specific parameters for the (potential) tourists' socio-economic characteristics, the 7 scenarios and the IVs of the lower-level nests.

Given the above model specification, the parameter for the IVs of the lower-level nests exceeded one, thus indicating the Nested Logit model to violate the model assumptions. It was therefore decided to model the two choice-stages of the tourist time allocation process using separate sequential MNL-models, where the participation choice model includes constants (base utilities) rather than the IVs for the underlying nests. The statistics of the four models are presented in Table 7.3, and Tables 7.4 through to 7.7 present the parameter estimates for these models. Empty cells in these tables indicate the exclusion of non-significant parameters. According to the likelihood ratio statistics, the estimated models all performed significantly better than the null-model. Also, the Rho-squares indicate the models to perform reasonably well (the differences between the participation and the trip choice models can be explained by the fact that the latter models are discrete choice models based on allocation scores, whereas the first is based on single choice data).

Table 7.3 Statistics for the sequential MNL-models for time allocation to tourist trips by adults

	Participation choice	Day-trips only	Holidays only	Both DTs and holidays
N° of Choice Sets	2524	226	483	1501
N° of Alternatives	4	2	6	7
Type of data	choice	◄	allocation	►
-2 [LL (0) - LL (B)]	2039.8603	79322.80	526831.81	1530955.1
N° of Parameters	27	18	61	88
McFadden's Rho (AIC)	.283775	.925471	.952299	.927338
Results in Table	7.4	7.5	7.6	7.7

The next two sections describe how socio-economic characteristics affect time allocation to tourist trips and how changes in the tourists' decision-making context induce people to re-allocate their time. First, however, it is noted that in the participation model, the constant of "day-trips only" is negative (-1.05), indicating that relative to the "no tourist trips"-alternative, the base-utility of this alternative is lower. Furthermore, the constant of "holidays only" did not prove to be significant, and the constant of "both day-trips and holidays" is positive (1.13). These constants more or less follow the distribution of observations at the aggregate level (see Table 7.1). Alternatively, the constants of the trips in the trip models are all negative. This means that, at the aggregate level, the majority of days are allocated to the reference alternative, i.e. non-tourist activities.

THE EFFECTS OF SOCIO-DEMOGRAPHIC CHARACTERISTICS ON TIME ALLOCATION TO TOURIST TRIPS BY ADULTS

As expected (see Table 7.4), *low income* is associated with lower participation rates in tourism because, relative to the highest income level, this level has a negative effect on the participation in holidays (-1.26) and day-trips and holidays (-1.33). In the trip models (see Tables 7.5 through to 7.7) it has neutral (parameter does not differ significantly from zero) or negative effects on all tourist trips. The *intermediate income* level has similar effects on the participation and trip choices. For this income level, however, the parameters are often less negative than for the lowest income level (except for the parameter for day-trips in the "day-trips only"-model and for extended holidays in the "both day-trips and holidays"-model).

Table 7.4 Participation choice model (alternative specific parameters; reference: "No tourist trips"); *t*-values in brackets

	Day-trips only	Holidays only	Both Day-trips and Holidays
Constant	-1.05 (-5.66)		1.13 (5.75)
socio-demographic characteristics			
Income < NLG 35,000		-1.26 (-8.69)	-1.33 (-8.63)
Income NLG 35-55,000		-.680 (-4.52)	-.519 (-3.56)
Education low			-.977 (-8.04)
Education intermediate			-.231 (-2.05)
Age <= 30 years			.296 (1.98)
Age 31-55 years			.275 (1.99)
Female			.230 (2.44)
Household size			
Child in household			
Urban residence			
At least one car	.405 (1.98)	.497(3.42)	.517 (3.31)
Permanent accommodation		1.29 (3.87)	1.19 (3.79)
Accommodation without place	1.10 (4.17)	1.66 (7.16)	1.97 (8.94)
Work: employed		.692 (4.90)	.707 (5.00)
Working hours per week	.017 (3.58)		
scenarios			
A: 12 extra free days per year		.631 (3.46)	
B: 12 free days less per year	-.638 (-1.97)		-.441 (-2.61)
E: free Fridays			
F: partner 12 extra free days per year			
G: partner 12 free days less per year			-.555 (-2.90)
I: improved student transport ticket	1.52 (4.55)		
J: older people +15% income			

Similar effects are also found for the three *educational levels* on the probability of selecting the participation category “both day-trips and holidays” and the allocation of time to the tourist trips. Exceptions include the positive parameter for the intermediate educational level in the “day-trips only”-model, and the (slightly) less negative effects of the lowest educational level on medium long holidays in the “holidays only”-model, and on nearly all trip types in the “both day-trips and holidays”-model. Apparently, lower educated people are much less likely to pursue both day-trips and holidays, but those that do choose to participate in these activities, do so more frequently than their higher educated counterparts.

Table 7.5 Trip choice model: “Day-trips only” (reference: “Non-tourist activities”); *t*-values in brackets

Constant	-2.36 (-32.5)
socio-demographic characteristics	
Income < NLG 35,000	
Income NLG 35-55,000	-.179 (-5.19)
Education low	-.460 (-9.51)
Education intermediate	.308 (6.63)
Age <= 30 years	
Age 31-55 years	-.142 (-3.89)
Female	-.373 (-10.9)
Household size	.170 (9.26)
Child in household	-.574 (-12.1)
Urban residence	-.152 (-4.15)
At least one car	-.110 (-2.61)
Permanent accommodation	-.461 (-4.76)
Accommodation without place	
Work: employed	-.258 (-6.98)
Working hours per week	
scenarios	
A: 12 extra free days per year	.184 (3.13)
B: 12 free days less per year	-.189 (-2.62)
E: free Fridays	.192 (2.17)
F: partner 12 extra free days per year	-.256 (-3.09)
G: partner 12 free days less per year	.170 (2.12)
I: improved student transport ticket	-.284 (4.24)
J: older people +15% income	

Table 7.6 Trip choice model: “Holidays only” (parameters for the numbers of days allocated to Short Breaks (SB), Medium Long Holidays (MLH), Extended Holidays (EH), Long Holidays (LH) and Extra Long Holidays (ELH); reference: “Non-tourist activities”); *t*-values in brackets

	SB	MLH	EH	LH	ELH
Constant	-4.45 (-32.4)	-3.87 (-57.4)	-3.76 (-50.8)	-4.06 (-47.3)	-4.19 (-28.8)
socio-demographic characteristics					
Income < NLG 35,000		-.439 (-6.55)		-.286 (-5.20)	
Income NLG 35-55,000		-.126 (-2.54)		-.248 (-5.68)	
Education low	-.881 (-14.3)	-.397 (-7.14)	-.193 (-4.80)		-1.35 (-14.2)
Education intermediate		-.424 (-7.52)			-1.17 (-11.9)
Age <= 30 years		-.244 (-3.91)		-.587 (-8.37)	-1.09 (-7.84)
Age 31-55 years	.654 (8.71)			-.251 (-4.12)	
Female	-.244(-3.84)		-.196 (-4.90)		
Household size	-.197 (-5.18)		-.122 (-6.99)		-.346 (-5.63)
Child in household	.412 (4.55)			.114 (2.41)	-.533 (-3.55)
Urban residence	-.555 (-9.07)				.331 (4.30)
At least one car	.460 (4.23)		.192 (3.06)	.390 (5.47)	
Permanent accommodation	1.77 (31.4)	.453 (6.52)		.543 (9.18)	
Accommodation without place		.164 (3.66)		.684 (17.5)	1.43 (17.5)
Work: employed	-.272 (-3.10)				-.429 (-4.96)
Working hours per week	-.012 (-4.34)	-.007 (-4.95)		-.012 (-8.91)	
scenarios					
A: 12 extra free days per year		.399 (6.00)		-.587 (-6.87)	.759 (6.80)
B: 12 free days less per year	-.284 (-2.86)				-1.11 (-5.70)
E: free Fridays	.435 (3.97)				
F: partner 12 extra free days per year	.420 (4.74)	.250 (3.40)		-.335 (-4.46)	.937 (-7.67)
G: partner 12 free days less per year	-.249 (-2.16)	-.265 (-2.61)	.430 (5.80)	-.158 (-2.17)	
I: improved student transport ticket	.693 (2.58)	.884 (5.19)			
J: older people +15% income			-.342 (-3.74)		

Older people (56 years and older) are less likely to participate in both day-trips and holidays because, relative to this group, the other two age categories have positive effects on this participation choice alternative (.296 and .275). Within the trip models the parameters indicate that, relative to the elderly, *young adults* (16-30 years) who have decided to participate in both day-trips and holidays are more likely to allocate time to day-trips (.443) and short breaks (.354) while the effects on longer holidays are neutral or negative (-.317, -.075, no significant effect and -.540). Within the other two trip models, there are no significant effects on day-trips, short breaks and extended holidays while the parameters for medium long, long and extra long holidays are again (strongly) negative (-.244, -.587 and -1.09). Those *aged between 31 and 55 years* also have neutral or negative parameters for longer overnight tourist trips in the trip models, while they have no, or positive effects on day-trips and short-breaks except for a small negative effect on day-trips in the "day-trips only"-model (-.142). Based on these parameters it can be concluded that older people are less likely to participate in both day-trips and holidays, but once they have decided to pursue tourist trips, they are more inclined to pursue longer holidays while younger people are often more likely to allocate time to day-trips and short breaks. These patterns are probably explained best by the fact that a

Table 7.7 Trip choice model: "Both Day-Trips and Holidays" (parameters for the numbers of days allocated to Day-Trips (DT), Short Breaks (SB) and Medium Long Holidays (MLH); reference: "Non-tourist activities"); *t*-values in brackets

	DT	SB	MLH
constant	-2.33 (-85.2)	-4.65 (-97.4)	-3.53 (-79.9)
socio-demographic characteristics			
Income < NLG 35,000	-.170 (-9.60)		-.339 (-8.97)
Income NLG 35-55,000	-.039 (-2.98)		-.131 (-4.99)
Education low	-.414 (-24.3)		-.168 (-5.07)
Education intermediate	-.284 (-21.1)	-.066 (-2.45)	-.100 (-3.59)
Age <= 30 years	.443 (18.2)	.354 (7.30)	-.317 (-7.12)
Age 31-55 years	.187 (8.02)	.339 (7.95)	-.230 (-5.85)
Female	-.190 (-15.1)		
Household size	-.082 (-11.6)	-.154 (-14.0)	-.079 (-7.62)
Child in household	-.206 (-10.8)		
Urban residence	.132 (11.2)	.215 (8.08)	
At least one car			
Permanent accommodation		1.25 (36.2)	.472 (12.7)
Accommodation without place	.276 (22.9)	.169 (6.41)	.153 (6.37)
Work: employed	-.191 (-11.8)		
Working hours per week			-.006 (-6.97)
scenarios			
A: 12 extra free days per year	.104 (4.67)	.403 (8.51)	.254 (5.58)
B: 12 free days less per year			
E: free Fridays	.076 (2.62)	.724 (14.3)	
F: partner 12 extra free days per year	-.075 (-2.89)	.278 (5.50)	
G: partner 12 free days less per year	.115 (4.14)	.194 (3.27)	
I: improved student transport ticket	.534 (15.9)	.671 (8.73)	.347 (4.35)
J: older people +15% income	.225 (7.01)		.153 (2.78)

proportion of the elderly is incapable to travel due to health issues, while those that are still healthy have enough time (and money?) to go on longer holidays.

According to the model parameters, being a *woman* has a small positive effect on participating in both day-trips and holidays (.230). However, within the trip choice models the parameters for females are either not significant or negative. In other words, women are more likely to participate in both day-trips and holidays, but once they have decided to pursue tourist trips, they allocate equal or less time to tourist trips than men. Perhaps these time allocation patterns should be attributed to the combination of both working and caring activities of women which often results in rather dispersed intervals of leisure time.

With regard to the household composition and location, the household size, living in an urban area and the presence of children do not have significant effects on the participation choices. Within the trip models, however, the *household size* has no or negative effects on the time allocated to the tourist trips except for a

Table 7.7 (continued) Trip choice model: "Both Day-Trips and Holidays" (parameters for the numbers of days allocated to Extended Holidays (EH), Long Holidays (LH) and Extra Long Holidays (ELH); reference: "Non-tourist activities"); *t*-values in brackets

	EH	LH	ELH
constant	-3.85 (-78.8)	-3.74 (-97.0)	-4.28 (-45.5)
socio-demographic characteristics			
Income < NLG 35,000	-.075 (-2.38)	-.574 (-17.1)	-.644 (-8.15)
Income NLG 35-55,000	-.108 (-4.66)	-.204 (-9.11)	-.395 (-7.71)
Education low		-.188 (-6.50)	
Education intermediate	-.114 (-5.26)	-.277 (-11.7)	-.937 (-15.0)
Age <= 30 years	-.075 (-2.02)		-.540 (-6.91)
Age 31-55 years	-.108 (-3.20)		
Female			-.337 (-6.54)
Household size		-.084 (-10.2)	-.123 (-3.92)
Child in household	-.276 (-11.3)		-.791 (-9.16)
Urban residence	.087 (4.06)	.273 (13.2)	.601 (12.2)
At least one car	.146 (4.15)		
Permanent accommodation	.399 (11.4)	.218 (6.01)	.653 (10.3)
Accommodation without place		.713 (33.2)	.937 (18.4)
Work: employed			-.309 (-4.11)
Working hours per week		-.003 (-5.38)	-.034 (-14.8)
scenarios			
A: 12 extra free days p.y.		.215 (5.70)	.474 (4.65)
B: 12 free days less p.y.		-.100 (-2.52)	-.991 (-5.26)
E: free Fridays	.198 (3.95)		
F: partner 12 extra free days	.107 (2.61)	.175 (4.76)	.204 (2.26)
G: partner 12 free days less			
I: improved student transp. ticket			
J: older people +15% income	.344 (6.91)		

positive effect on the number of day trips in the “day-trips only” model (.170). The *presence of children* strongly impedes extra long holidays (-.533 and -.791), and its effect on the other tourist trips varies depending on the model. Living in an *urban area*, finally often has neutral or positive effects on the time allocated to tourist trips except for day-trips in the “day-trips only”-model (-.152) and short breaks in the “holidays only”-model (-.555).

With regard to accommodation and transport opportunities, the possession of a *car* and the *possession of a tourist accommodation (with or without a permanent location)* both have very strong positive effects on all participation choice categories. The only exception is the parameter for permanent accommodations and “day-trips only”, that did not prove to be significant. Within the trip models, *car ownership* has no or only small positive effects on the trip choices except for day-trips in the “day-trips only” model (-.110). Apparently, the lack of a car does refrain people from allocating (more) time to holidays, while it does not constrain day-tripping. Perhaps these observations are related to the fact that people often have a lot of luggage when they go on holidays, while these problems are less substantial for day-trips.

The possession of a *tourist accommodation with a permanent location* has a strong positive effect on all overnight trips (except for extended and extra long holidays in the “holidays only”-model), in particular on short breaks (1.77 and 1.25), and on extra long holidays in the “both day-trips and holidays”-model. The effect of the permanent accommodation on day-trips is either not significant (in the “both day-trips and holidays”-model) or strongly negative (-.461 in the “day-trips only”-model). In contrast, the effect of the possession of a *tourist accommodation without a (permanent) location* is always positive (or non-significant), especially on extra long holidays (1.43 and .937). These parameters clearly demonstrate that people who have decided to purchase a personal accommodation are very likely to allocate a lot of time to tourist trips, extra long holidays in particular. People who have purchased tents, boots, caravans and (second) houses with a permanent location are also more inclined to allocate a lot of time to short breaks.

Work, finally, has a positive effect on participation in holidays (with (.692) or without (.707) day-trips), whereas the probability of participating in day-trips only increases with the number of working hours (.017). Within the trip models, however, being employed and the number of working hours both have either neutral or negative effects on the trip types, extra long holidays in particular. These parameters clearly demonstrate the effect of time constraints due to working activities. Given the negative effects of lower income levels, and an (assumed) positive relationship between the employment situation and income levels, it can be concluded that being employed impedes participation in tourist trips for lower income groups only.

TOURIST TIME RE-ALLOCATION: THE EFFECTS OF CHANGES IN THE TOURISTS' DECISION-MAKING CONTEXT ON TIME ALLOCATION TO TOURIST TRIPS

In addition to the effects of the socio-demographic characteristics, the sequential models also comprise the effects of several changes in the tourist's decision making context. The discussion first focuses on the scenarios that did not have (significant) effects on the participation choice level, i.e. the 4-day working week with free Fridays (scenario E), the extra free days for the partner (scenario F) and the increase of the household income for the retired (scenario J). Next, scenarios that did bring about changes at the participation choice level are discussed (i.e., scenarios A, B, G and I).

With regard to altered working conditions, the 4-day working week with free Fridays (scenario E) and the increase in the number of free days for the partner (scenario F) generally have positive effects on the time allocated to the various trips. As expected, *free Fridays* have positive effects on the time allocated to shorter trips including day-trips (.192 and .076), short breaks (.435 and .724) and extended holidays (.198 in the "both day-trips and holidays"-model). The *12 extra free days for the partner* have negative effects on the time allocation to day-trips (-.256 and -.075) and on long holidays in the "holidays only" model (-.335). Apparently, in addition to simply increasing the number of trips, the latter scenario also induces people to increase the duration of their tourist trips or substitute them with (more) shorter trips.

The *increase in income for pensioners* (scenario J), has small positive effects on day-trips (.225), medium long (.153) and extended holidays (.344) in the "both day-trips and holidays"-model. Strikingly, the only effect of this scenarios in the "holidays only"-model is negative (-.342 on extended holidays). Since other (positive) effects do not compensate for this effect, the increase in income for pensioners apparently has a negative effect on time allocation to tourist trips for this segment.

Four of the examined changes in the tourists' decision-making context proved to have effects on the participation choices. Granting people *12 extra free days* (scenario A), for instance, has a positive effect on the probability of participating in "holidays only" (.631). At the trip choice level, this scenario has neutral or positive effects on all tourist trips except for long holidays in the "holidays only"-model (-.587). Since the latter effect is compensated by other positive effects (in other words, the longer holidays are substituted by either extra long or (more) medium long holidays), it can be concluded that granting people 12 extra free days will generally increase the time allocated to tourist trips.

In contrast, *decreasing the numbers of free days* (scenario B), has a negative effect on the "day-trips only" (-.638) and on the "both day-trips and holidays" (-.441) participation alternatives. In the trip models, this scenario has negative effects on the time allocated to day-trips in the "day-trips only"-model (-.189), on short breaks (-.284) and extra long holidays (-1.11) in the "holidays only"-model, and on long (-.100) and extra long (-.991) holidays in the "both day-trips and holidays"-model. The analogous scenario *for the partner* (G) only shows a negative effect on the participation in both day-trips and holidays (-.555). In the "holidays only"-model, this scenario has negative effects on short breaks (-.240), medium long (-.265) and long (-.158) holidays and a positive effect on extended holidays (.430). Hence, in this model people indicate to substitute trips of different length in response to the partner's changed working conditions, while the total amount of time allocated to tourist trips slightly decreases. In the other models, this scenario only has positive effects, in particular on short trips: day-trips (.170 and .115) and short breaks (.194). Perhaps the partner's increased number of working days are made up for by using the extra income for tourist trips that can be pursued during the weekends.

The final scenario that had an effect on both the participation and the trip choices *improved* the *conditions of the public transport ticket for students* (scenario I). This scenario has a rather strong positive effect on the "day-trips only"-alternative (1.52), as well as positive effects on nearly all shorter trips (up to 9 days) in the three trip models. Strikingly, the effect of this scenarios in the "day-trips only"-model is negative (-.284). This observation may possibly be explained by the fact that people who previously did not allocate time to tourist trips were induced to change their participation choice and allocate some time to day-trips. If the average amount of time allocated to day-trips by these "new tourist" is lower than that by the students that already pursued day-trips before the change in the decision making context, at the aggregate level this scenario indeed has a negative effect on day-trips in the "day-trips only"-model. However, the amount of time allocated to day-trips at the aggregate level may still increase because of the increased participation rates.

7.4 The "child"-model of tourist time allocation behaviour

As mentioned before, the general model for tourist time allocation processes for day- and overnight trips (and resource re-allocation behaviour) cannot be applied to those aged 0 to 15 years due to the lack of relevant data. Consequently, with regard to the tourist trip patterns of children, the *MERLIN*-system is reduced to overnight trips and also excludes resource re-allocation behaviour. The model for tourist time allocation processes for this group is presented in Figure 7.2.

This section discusses the overall time allocation to holidays by children and the estimation results of the models for this group.

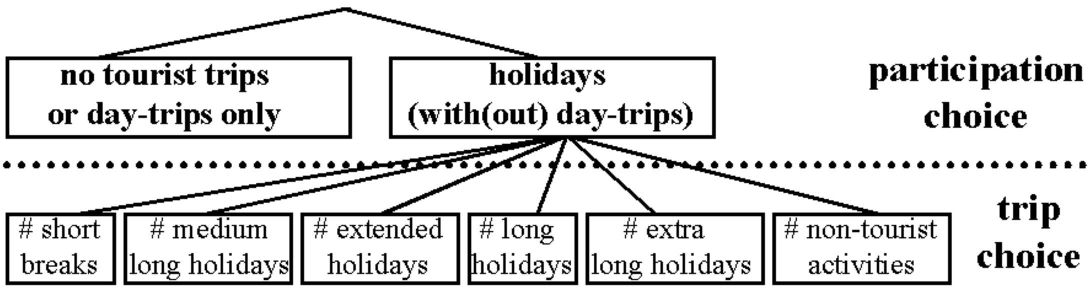


Figure 7.2 The model for tourist time allocation processes for day- and overnight trips for children under 16 (# : the time (expressed in the number of days) allocated to a particular trip)

7.4.1 Overall time allocation to tourist overnight trips by children

The 1998 CVO-data included 726 children between 0 and 15 years, 612 of which pursued a total of 1313 overnight tourist trips in 1998 (2.14 holidays per participant). Based on the exact duration of their holidays (in days), these 1313 holidays were classified according to the 5 overnight tourist trip categories (short-breaks and medium long, extended, long and extra long holidays). Finally, in order to even out the two models for tourist time allocation, the time that these children had allocated to the various trips was calculated using the same conversions as outlined previously for the "adult"-model. Similarly, the time allocated to non-tourist trips was calculated by deducting the numbers of days spent on overnight tourist trips from 365 days. Table 7.8 displays the average number of days allocated to tourist trips in 1998 for the two "child"-participation groups. Both the exact (based on the CVO-data) and the approximated (following the "adult"-conversion) number of days allocated to overnight tourist trips are reported, indicating that the converted data slightly overestimate the time allocated to tourism. This observation is confirmed by regressing the exact number of days (independent) onto the approximated numbers (dependent): the intercept differes slightly, but not significantly from zero (.234; t = .882), and the slope exactly equals one (1.0oo; t-value with respect to (the deviation from) unity = .0000; R = .948; (adjusted) R-square = .898).

Table 7.8 Approximated number of days allocated to overnight tourist trips in 1998 per participation group for children

	No tourist trips or day-trips only	Holidays (with(out) day-trips)	Total
1998-CVO-data (exact)	0.00	19.66	16.57
after conversion	0.00	19.84	16.72
(N)	(114)	(612)	(726)

On average, children allocate almost 17 days per year to overnight tourist trips (see Table 7.8); for those who actually pursue these trips, the time allocated to holidays amounts to 20 days per year. The latter number resembles the time allocated to overnight trips by adults who only pursue holidays. Finally, it should be noted that the holiday participation rate of children exceeds that of adults (83% vs. 79%).

7.4.2 The tourist time allocation models for children (0-15 years)

Unfortunately, for the "child"-model too, simultaneous estimation of the joint participation-trip model was not possible. Similarly, the parameter estimate for the Inclusive Value in the simultaneous model exceeded one, thus violating the model assumptions. Like the "adult"-model, therefore, the two-staged process of time allocation to tourist trips for children is modelled using two separate MNL-models, where the participation choice model includes a constant for the "Holidays (with(out) day-trips)"-alternative (using "No tourist trips" as a reference) rather than the IVs of each individual for the underlying nest. The socio-economic characteristics that were entered into the two "child"-models are listed in Table 7.9. The time allocated to non-tourist activities was selected as the base alternative in the trip model and all socio-economic variables and constants were entered as alternative specific parameters. Parameters that were not significant at the 95% confidence level were excluded. The statistics of the models are displayed in Table 7.10, and the parameter estimates are presented in the Tables 7.11 and 7.12.

According to the likelihood ratio statistics, the estimated models both performed significantly better than the null-model. Also, the rho-squares indicate the models to perform rather well (the differences between the participation choice model and the trip choice model can again be explained by the fact that the latter model is a discrete choice model based on allocation scores, whereas the first is based on single choice data).

Table 7.9 The socio-demographic variables that were used to explain time allocation choices of children [cont. = continuous]

Variable	Levels	Coding	N/average
Net Household Income	less than NLG 35,000	[1,0]	164
	NLG 35,000-55,000	[0,1]	369
	more than NLG 55,000	[0,0]	193
Age	0-1 years	[1,0,0]	75
	2-5 years	[0,1,0]	194
	6-13 years	[0,0,1]	332
	13+ years	[0,0,0]	125
Gender	female	[1]	347
	male	[0]	379
No. of People in Household		[cont.]	4.23
Residence	urban	[1]	221
	rural	[0]	505
(a) Car(s) in household	yes	[1]	615
	no	[0]	111
Accommodation with permanent place	yes	[1]	40
	no	[0]	686
Accommodation without place	yes	[1]	305
	no	[0]	421
Possession of Skis	Ski	[1]	58
	no ski	[0]	668
Holiday region	north	[1,0]	261
	mid	[0,1]	211
	south	[0,0]	254
Social Class	Range 1-5	[cont.]	2.74

THE EFFECTS OF SOCIO-DEMOGRAPHIC CHARACTERISTICS ON TIME ALLOCATION TO TOURIST TRIPS BY CHILDREN

Only four socio-demographic variables significantly affect the participation choices of children (see Table 7.11). First, relative to the oldest *age group* (13-15 years), being a baby (0-1 years), has a negative effect on the probability of participating in overnight tourist trips (-.608). At the trip level, this variable also has a negative effect on time allocation to holidays, in particular on extended (-.300) and long holidays (-.827). The other age levels do not have effects at the participation choice level, and only affect the time allocation to holidays. Relative to the oldest children (13-15 years), being 2 to 5 years old has negative effects on medium long and long holidays (-.148 and -.217), while it has positive effects on extended (.138) and in particular extra long holidays (.764). Similarly, being 6-12 years old has a negative effect on long holidays (-.271), while it has a positive effect on extended (.314) and again extra long holidays (1.06). Apparently, babies and children in secondary school are less likely to go on extra long holidays than toddlers and children in primary school.

Table 7.10 Statistics for the sequential MNL-models for time allocation to tourist trips by children (0-15 years)

	Participation choice model	Holidays (with(out) day-trips) model
N° of Choice Sets	726	612
N° of Alternatives	2	6
Type of data	choice	allocation
-2 [LL (0) - LL (B)]	422.22058	674820.34
N° of Parameters	5	52
McFadden's Rho (AIC)	.409579	.959261
Results in Table	7.11	7.12

The second and third variables to affect the children's participation choices are the possession of a *tourist accommodation with a permanent place and without a permanent place*. Both have strong positive effects for children on participating in overnight holidays (2.21 respectively .958). At the trip choice level, the possession of a tourist accommodation (with or without a permanent place) has neutral or positive effects on all holiday types except for medium long (-.097) and extended (-.195) holidays in case of non-permanent accommodations. The strong positive effect of both accommodation types on long (.626 and .776) and extra long (2.28 and .801) holidays mirrors the effects of these tourist commodities in the adult models. Similarly, accommodations with permanent positions have strong positive effects on the time allocated to short breaks (1.88).

The fourth variable to affect both the participation and the trip choices is *social class*. Given the coding of this continuous variable (1 = highest social class), children from lower classes are less likely to participate in overnight tourist trips (-.366). At the trip choice level, however, lower classes have positive effects on the time allocated to medium long (.097) and extra long (.177) holidays, while they have a negative effect on extended holidays (-.121). Apparently, children from lower social classes are less often inclined to participate in tourist trips, but once the decision to participate has been taken, they allocate (slightly) more time to these trips than children from higher social classes.

Table 7.11 Participation choice model for children (alternative specific parameters; reference: “No tourist trips or day-trips only”); *t*-values in brackets

Constant	2.43 (7.33)
socio-demographic characteristics	
Income < NLG 35,000	
Income NLG 35-55,000	
Age 0-1 years	-.608 (-2.00)
Age 2-5 years	
Age 6-12 years	
Female	
Household size	
Urban residence	
At least one car	
Permanent accommodation	2.21 (2.16)
Accommodation without place	.958 (4.02)
In possession of ski	
Holiday region: North	
Holiday region: Mid	
Social class	-.366 (-3.72)

The remaining socio-demographic characteristics only affect the children’s trip choices. Relative to the highest *income level*, for instance, intermediate income is negatively related to medium long (-.097), long (-.196) and extra long (-.401) holidays. In contrast, lower levels of income have a strong negative effect on long holidays (-.646), but an equally strong positive effect on extra long holidays (.666). Perhaps the latter effect can be explained by constraints related to the working schedules of the more affluent parents.

Another variable that exerts an influence on the trip choices of children is *gender*: girls are associated with decreasing numbers of days allocated to short breaks (-.189).

With regard to household composition and location, the *number of people in the household* is negatively related to medium long holidays (-.103) and positively related to extended (.051) and extra long (.115) holidays. The later observations oppose the effect of household size in the “adult”-model. Living in an *urban area* has positive effects on three of the holiday types, i.e. short breaks (.158), long (.136) and extra long (.489) holidays. With regard to the *holiday region* it can be concluded that, relative to living in the south, Dutch children living “above the rivers” (in the north and middle parts of the country) have positive effects on extended (.105 and .114) and extra long holidays (1.07 and 1.36), while they have negative effects on medium long holidays (-.221 and -.130).

Table 7.12 Trip choice model for children under 16 (parameters for the numbers of days allocated to Short Breaks (SB), Medium Long Holidays (MLH), Extended Holidays (EH), Long Holidays (LH) and Extra Long Holidays (ELH); reference: "Non-tourist activities"); *t*-values in brackets

	SB	MLH	EH	LH	ELH
Constant	-5.65 (-59.0)	-3.54 (-31.1)	-4.08 (-35.8)	-4.34 (-61.8)	-9.46 (-28.8)
socio-demographic characteristics					
Income < NLG 35,000				-.646 (-12.4)	.666 (4.49)
Income NLG 35-55,000		-.097 (-2.41)		-.196 (-5.91)	-.401 (-3.18)
Age 0-1 years			-.300 (-3.31)	-.827 (-10.9)	
Age 2-5 years		-.148 (-3.18)	.138 (2.29)	-.217 (-4.94)	.764 (4.90)
Age 6-12 years			.314 (5.77)	-.271 (-6.88)	1.06 (7.31)
Female	-.189 (-3.43)				
Household size		-.103 (-4.63)	.051 (2.74)		.115 (2.37)
Urban residence	.158 (2.78)			.136 (4.20)	.489 (5.05)
At least one car	.225 (2.43)		-.221 (-4.62)	.438 (7.61)	-.357 (-2.89)
Permanent accommodation	1.88 (30.6)	.488 (7.16)	.197 (2.81)	.626 (12.1)	2.28 (22.5)
Accommodation without place	.385 (6.94)	-.097 (-2.36)	-.195 (-5.19)	.776 (23.7)	.801 (8.30)
In possession of ski	-.300 (-2.81)	.195 (2.99)		.182 (3.81)	
Holiday region: North		-.221 (-4.60)	.105 (2.39)		1.07 (7.13)
Holiday region: Mid		-.130 (-2.64)	.114 (2.51)		1.36 (9.23)
Social class		.097 (-5.32)	-.121 (-7.54)		.177 (3.65)

Finally, with regard to transport and tourist opportunities, the availability of (a) *car(s) in the household* has positive effects on short breaks (.225) and long holidays (.438), while it has negative effects on extended (-.221) and extra long holidays (-.357). Children owning *skis*, finally, are associated with higher numbers of days allocated to medium long (.195) and long (.182) holidays and lower numbers of days allocated to short breaks (-.300).

7.5 Conclusion and discussion

The modelling approach discussed in this chapter aimed at describing the relationship between different tourist trips over a one-year period in terms of the time allocated to these trips. In contrast to the continuous micro-economic time-allocation models, tourist time allocation choices are viewed here as a two-stage process. First, the (potential) tourist decides whether or not to participate in day-and/or overnight tourist trips - a discrete choice. Next, conditional on this participation decision, the individual allocates a certain amount of time to the various tourist trips - a more continuous allocation choice process. The latter choice is also modelled using a discrete choice model because it can also be viewed as a repeated choice to allocate one day to either a tourist or a non-tourist activity. By using sets of Multinomial Logit models, tourist time allocation processes were modelled in accordance with these two stages in the decision making process. Two empirical sets of models - one for children and one for adults - were calibrated and discussed.

The proposed time-allocation model is capable of representing the overall amount of time allocated to tourist trips as well as the role of each trip type within the annual set of tourist trips. Given the available data, the proposed "child"-model included various types of holidays, while the "adult"-model also comprised the relationships between day-trips and these holidays - trips that up till now have been modelled independently[23]. The "adult"-data demonstrated that people who make only day-trips or only overnight trips often allocate less time to tourist trips, whereas people who participate in both day-trips and holidays on average spend twice as much time on tourist trips (about 44 days). Strikingly, the time allocated to overnight tourist trips by children almost coincided with that allocated by adults in the "holidays only" nest.

The choices between the participation alternatives and the subsequent allocation of time to the various day- and overnight tourist trips were related to socio-economic variables. Overall, the model parameters in both the "adult"- and the "child"-model confirmed the generally known relationships between socio-demographics on the one hand, and participation in tourism on the other. However, several parameters did not have the expected sign or size. Some of these abnormalities could be explained, but others can only be attributed to chance.

In addition to representing the present situation, the "adult"-model also incorporated parameters that indicate the effects of changes in the traveller's financial, time and transport circumstances. Seven scenarios were examined in terms of their effects on time re-allocation. Increasing the respondent's number of free days, decreasing the respondent's or his or her partner's number of free days and improving the conditions of the student public transport ticket were shown to have an impact on both the participation in tourist trips and the time allocated to these trips. The other scenarios, including the 4-day working week with free

[23] Due to data constraints, the child-model did not include information regarding day-trip behaviour of children. This is rather unfortunate because children are an important target group for this part of the tourism industry.

Fridays, the extra free days for the partner, and the increase of the household income for the retired, only proved to affect the latter choice processes only.

Given the annual set of tourist trips (including the duration of each trip) described in this chapter, *MERLIN* will continue the (simulation of the) tourist decision making process by further profiling the generated overnight tourist trips. The next chapter addresses the choice of travel companions.

8 The Choice of Travel Party

This is the first chapter in a series that describe the results of the application of a CHAID-based algorithm to induce decision-making rules for each facet of the tourist decision-making process. With regard to the induction, presentation and validation of decision tables (DTs), this chapter formally describes the approach and the statistics and confusion matrix that are presented for each DT. With regard to tourist decision-making, this chapter considers the choice of travel party.

8.1 Introduction

The previous chapter discussed tourist time allocation processes to generate annual tourist trip programs. The next step is to select the travel party for each (overnight) trip in this program. As argued in chapter 3, this, and all subsequent profiling decisions for tourist trips will be modelled using a rule-based approach. More specifically, a CHAID-based algorithm is used to induce exclusive and exhaustive sets of decision rules that are laid down in decision tables (DTs). Each chapter describing one or more of these DTs will have a similar structure. First, the facet of the tourist decision making process under consideration is made operational. This comprises the definition of the choice variable, i.e., the alternatives in the choice set $\S_N$ (where N denotes the facet under consideration). If the number of alternatives in the choice set is rather large, this may also come down to the definition of a hierarchy of choice sets because this better matches non-compensatory decisions strategies[24] (Payne, 1976). In these cases, sets of alternatives are systematically excluded to eventually arrive at the preferred choice. Alternatively, if the choice facet comprises multiple dimensions, the facet may be described using two or more non-hierarchical DTs. Timing choices in chapter 9, for example, include the choice of season as well as the decision whether or not to travel during a school holiday period.

Given the structure of the choice set(s), the condition variables that are entered into the CHAID-based algorithm are discussed as well as the settings of the algorithm. Next, the structure of the generated DTs is presented, including a discussion of the most notable decision rules. Finally, several statistics of the DT are considered, and the validity of the DT is assessed by means of a confusion matrix that presents the probabilities of predicting the *i'*-th alternative while the *i*-th alternatives was observed.

---

[24] From a methodological point of view, the use of a hierarchy of choice sets may be necessary when the response distribution of the population over the alternatives is rather unbalanced. In this case, the CHAID-based algorithm may not be able to identify decision rules that identify the less popular choice alternatives.

As the first in the series of chapters describing the profiling decisions for tourist trips, this chapter discusses at length the approaches to deriving and validating DTs. First, therefore, the next section discusses the CHAID-based algorithm. Subsequently, the common structure for chapters is followed.

8.2 A CHAID-based approach to deriving tourist decision rules

Chi-squared AID (CHAID) is an offshoot of AID (Automatic Interaction Detection) and was designed by Kass (1980) for categorised (nominal and ordinal) variables[25]. The CHAID algorithm partitions the observed choices into mutually exclusive and exhaustive subgroups by maximising the significance of the chi-squared (χ^2) statistic at each partition (Kass, 1980). This way, it considers the whole distribution of the actions, not only their means and variances (Strambi & Van de Bilt, 1998). Also, it has a built-in significance testing with the consequence of using the most significant condition variable rather than the most explanatory when splitting and/or merging the population (Kass, 1980). The χ^2 of a DT is given by:

$$\chi^2 = \sum_{i=1,j=1}^{I,J} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

where i (i = 1, 2,.. I) denotes a choice alternative, j (j = 1, 2,.. J) denote a decision rule, O_{ij} is the frequency (i.e. the number of observations) of the i-th choice alternative and the j-th decision rule and E_{ij} is the *expected* frequency of the i-th choice alternative and the j-the decision rule under the assumption of independence between the conditions and the response distribution (for the actions). The expected frequency of a cell in a decision table is computed by multiplying the total number of observed cases of the j-the decision rule and the total number of observed cases of the i-th choice alternative, and dividing this by the overall number of observed cases. The degrees of freedom of the χ^2-statistic is given by Df = (I-1)(J-1).

Based on a decision table, the CHAID-based algorithm computes the χ^2-statistics and corresponding "pairwise" p-values for each pair of the condition states that is eligible to be merged. In case of categorical conditions (*free* or *flexible*

[25] Software tools also allow for continuous variables. For continuous dependent variables, an F-test (instead of the Pearson chi-squared or likelihood-ratio test) is used to identify segments in the population that differ significantly with regard to the dependent variable (SPSS, 1998). Strictly speaking, these techniques are not CHAID because they use different significance tests.

predictors in the original CHAID-analysis), any grouping of categories is permitted, whereas in case of ordinal condition variables (*monotonic predictors* in the original CHAID-analysis) only contiguous categories may be grouped together. In case of continuous condition variables, the quantitative scale is first divided into rages of values using an *equal frequency method*. These ranges are than treated as “normal” ordinal categories. In this process, missing values of ordinal condition variables are viewed as a separate condition state which can be grouped together with any of the other states (in case of missing values, monotonic predictors are referred to as *floating predictors* in the original CHAID-analysis).

If the largest of all pairwise *p*-values is greater than a specified α-level, this pair of states is merged into a single compound condition state, and the whole process is repeated (with this new compound state) until the largest pairwise *p*-value at a certain stage is smaller than the specified α-level. Finally, for each compound condition state consisting of three or more of the original states, the algorithm finds the most significant binary split. If the significance is beyond a critical value, this split is implemented and the algorithm returns to the merging stage. The merging process is repeated for each condition variable and its states in turn. Next, for each optimally merged condition variable, the (adjusted) *p*-value of the (reduced) decision table is computed. If there has been no reduction of the original table, a χ^2-test can be used that is conditional on the number of categories of the condition variable. However, if the table has been reduced, a more conservative significance test should be used to avoid the risk of capitalising on chance in searching for the optimal grouping of condition states. In the CHAID-based algorithm, the *adjusted p-value* is obtained by using a proper *Bonferonni multiplier*. Basically, this multiplier (*B*) is determined by the number of ways a condition variable *of a given type* with *c* original condition states can be reduced to *r* states ($1 \le r \le c$). For *monotonic* condition variables B_{mon} is:

$$B_{mon} = \left(\frac{c-1}{r-1} \right)$$

For *free* condition variables B_{free} is:

$$B_{free} = \sum_{i=0}^{r-1} (-1)^i \frac{(r-1)^c}{i!(r-1)!}$$

Finally, for *floating* condition variables B_{float} is:

$$B_{float} = \binom{c-2}{r-2} + r\binom{c-2}{r-2} = \frac{r-1+r(c-r)}{c-1} B_{mon}$$

Given the adjusted *p*-values of all conditions, the most significant condition variable (i.e., lowest *p*-value) is isolated, and if the *p*-value of this condition is less than or equal to the specified α-level, the group of observations is split according to the (merged) states of this condition variable. If no condition variable has a significant *p*-value, the group is not split, and the process is terminated. For each partition of the data that has not been analysed, the algorithm returns to the first step. The tree-growing process continues until all subgroups have either been analysed or contain too few observations.

The CHAID-algorithm may not find the optimal split for a condition variable because the merging process is terminated when all remaining states are found to be statistically different. "*Exhaustive CHAID*" is a modification of the original CHAID-analysis (Biggs *et al.*, 1991) that counteracts this bias by continuing to merge the states of the condition variables until only two states remain. The *p*-values and the corresponding sets of (compound) states at each successive stage are stored. Next, the successive merges for each condition are considered, and the set of (compound) states that gives the strongest association with the decision variable is selected as the optimal merge for each condition variable. Exhaustive CHAID than proceeds with computing the adjusted *p*-values and selecting the most significant condition. The DTs for the tourist trip profiling decisions in the MERCIN-system are derived using an algorithm based on the exhaustive CHAID-algorithm.

8.3 The question of whom to travel with

8.3.1 The choice set

Given the annual set of tourist trips, i.e. the tourist trip program, the next step of the MERCIN-system is the selection of the travel party. Due to the data limitations discussed in chapter 5, only the travel party (and all subsequent choice facets) of overnight tourist trips can be considered. The choice of travel companions is considered as the first profiling decision because in many cases the composition and size of the travel party determines the availability and attractiveness of the other facets of the tourist decision. The presence of schoolgoing children in the travel party, for instance, precludes holidays outside school holidays and/or weekends. The absence of children, on the other hand, may increase the attractiveness of long distance travelling, certain means of transport, and/or lodging facilities.

The travel party alternatives considered by the *MERLIN*-system are based on the presence of children (including their age) and the party size. The first dimension is included because, as argued before, this determines the availability and attractiveness of the choice alternatives of decisions yet to come. The latter dimension is introduced because very small (single travellers in particular) and very large travel parties may become increasingly important to the travel industry due to demographic developments like the increase of single person households and the greying of the population. Hence, *MERLIN* considers the following travel party options: (1) alone; (2) adults (20+ years) only; (3) with schoolgoing children (6-14 years); (4) with other children (0-5 or 15+ years); (5) party of 9 or more people (regardless of their ages); and (99) unknown. Since these categories are exclusive, the parties of adults only, and parties with children always include at least 2, and at most 8 members. If both schoolgoing and "other" children are part of the travel party, these parties are characterised as parties with schoolgoing children because these schoolgoing children are more likely to restrict other choice facets (e.g., the timing of the holiday).

The last category, "unknown", is dictated by data considerations. This is explained as follows. In order to reduce the task load for respondents, the CVO-quarterly measurements record detailed information on the two longest holidays of each quarter only. If people make three or more holidays in a particular quarter, only the most important information is recorded, and several other holiday-related variables, including, the composition of the travel party for the third (and subsequent) holiday(s) are missing in the data set. In this process, the recording order of the holidays in a quarter is based on the duration of the holidays, where longer holidays are recorded first. Chances are that the characteristics of these additional holidays in a quarter differ significantly from the main holidays because, due to the registration protocol, they are likely to be shorter. Also, it can be hypothesised that a large proportion of these additional holidays are made by people who own a tourist accommodation because it was shown in the previous chapter that this variable has a strong positive effect on overnight trip frequencies. It was therefore decided to derive two sets of decision rules for the choice of travel party, one with, and one without the "unknown" category. The effect of the inclusion of this category was assessed by validating the *MERLIN*-system with both sets of rules. Based on these validation results, it was decided to include the "unknown" category because without this category, the MERLIN-system seriously underestimated the number of domestic trips, and holidays to/with permanent tourist accommodations owned by the tourist.

8.3.2 The condition variables

Research on the social aspects of tourist decision-making often focuses on the influence of the (family) life-cycle on holiday choices and/or on group (family) decision making process (Dellaert *et al*., 1998b; Enneking, 1979; Jenkins, 1978; Davis & Rigaux, 1974). To our knowledge, the selection of the travel party itself has never been examined to date. As a consequence, little is known regarding the motives for selection certain travel partners. Based on the conceptual framework, the MERLIN-system assumes the selection of the travel party (and all other tourist choices) to be determined by the personal and household conditions (including both preferences and constraints), system and institutional conditions, and conditions resulting from precious decisions. Appendix 1 introduces the condition variables, including their states and coding, that are used. For the choice of travel party, only the *conditions common to all choice facets* (Table A1.1) are entered into the CHAID-based algorithm. The remainder of this section describes these conditions. Subsequent chapters will only discuss the conditions that are newly introduced and/or specific to that choice facet.

The conditions that are entered into the CHAID-based algorithm for the choice of travel party (and all subsequent choice facets) can be grouped into several categories. First, two conditions represent the *tourist trip program choices* of the tourist, including the number of day-trips and the number of holidays. Since the number day-trips is not known for children younger than 15 years, this variable comprises a category "younger than 16 years". For adults, the day-trip frequencies are categorised into 11 ordinal states to account for the fact that these frequencies are based on a rough estimate of the annual number of day-trips. In contrast, the annual number of holidays are assumed to be more accurate because the CVO-panel member had recorded their holidays throughout the year (see chapter 6); this condition variable is therefore entered as a continuous variable and, for each DT, the CHAID-based algorithm will categorise it using the equal frequency method.

Second, there are condition variables describing *decisions that have been taken previously regarding the trip under consideration*. Since the choice of travel party is the first profiling decision, only the duration conditions the planning process for a holiday at this stage.

Third and finally, there is a group of variables describing the *personal and household characteristics* of the traveller, including:

(a) *Socio-demographics*, such as age, income, educational level, civil status, presence of children in the household and household size;

(b) Several conditions representing constraints and opportunities arising from the *working situation*: having a (part-time of full-time) job or not, the exact number of working hours, (the frequency of) working during weekends, the number of free-days in 1998 (the sum of paid holidays, days off resulting from shorter working hours (Dutch: *ATV*- or *ADV-dagen*), and paid holidays passed on from previous years), and finally the presence of restrictions imposed on the timing of holidays;

(c) *School holiday region:* with regard to school holidays, the Netherlands are divided into three regions, north, mid and south. In order to reduce congestion problems and increase occupancy rates of facilities, the beginning and ending of school holidays in these regions are spread in time (usually at a 1-week interval);

(d) Several variables describing the *location of residence*, including the province of residence, and two variables describing the level of urbanisation of the city of residence. These conditions capture regional differences and possible differences between urbanites and country people; and finally,

(e) Four variables capturing *tourist opportunities*, including the possession of cars, skis, tourist accommodations (boat, tent, caravan) without a permanent place, and/or tourist accommodations (boat, tent, caravan, second house) with a permanent place.

The above conditions are assumed to affect tourists' choices of travel parties. To conclude this section on the condition variables, three notes are in place. First, system and institutional conditions are not included due to data limitations and/or because they equally apply to all holidays (e.g., opening hours of facilities, the possession of a driver's license or weather conditions). Secondly, for continuous condition variables, the "equal-frequency method" is applied to create categorical states. Third and finally, it is recalled that not all these condition variables will actually be included in the decision rules because the CHAID-based algorithm will only select those conditions that, at some point in the decision tree, contribute most significantly to identifying different segments of the population.

8.3.3 Stopping criteria of the CHAID-based algorithm

Given the choice set and the condition variables for the choice of travel party, several decisions regarding the stopping criteria of the CHAID-based algorithm remain. First, the α-level that controls the tests for splitting and merging of condition states has to be specified. In the original study by Kass (1980), the merging criterion was "significance at the 5% level", and the splitting criterion was 4.9%. The resulting contingency table was considered significant if the *p*-value did not exceed the 5%-level after correction by the proper Bonferonni multiplier. For the algorithm based on the exhaustive CHAID-algorithm that is used in this study, the α-level for category merging is irrelevant, because this algorithm continues to merge (compound) categories until only two categories remain, to ultimately select the set of categories with the smallest *p*-value. Hence, only a 5% significance level for predictor eligibility is specified.

Secondly, the minimum number of observations before and after splitting can be specified. If these numbers are set rather low in relation to the total number of observations in the data set, decision rules at lower levels of aggregation are obtained. This advantage, however, comes at the cost of more sizeable sets of rules and the risk of overfitting on the training data (i.e., low transferability to new cases). Stricter stopping criteria, on the other hand, may lead to rulebases that are unable to identify certain choice alternatives. Since the effects of stopping criteria heavily depend on the characteristics of the data, a sensitivity analysis is performed for each DT to determine the numbers of observations that balance these advantages and disadvantages. For each DT, several combinations of before/after stopping criteria (i.e., 35/15, 50/20, 60/25, 80/35 and 100/45) are assessed in terms of the predictive ability of the resulting DT. Three indicators are used:

(1) the ability of the DT to repredict the original observations as indicated by the *percentage of correctly resubstituted observations*;

(2) a 4-fold cross-validation[26] of the ability of the DT to repredict the original observations as indicated by the *percentage of correctly cross-validated observations*; this approach was selected based on a empirical comparison of a 4-fold and a 10-fold cross-over approach, and validation based on partitioning the data in a 75% training and a 25% validation set; and finally

(3) the DT's ability to reproduce the observed choice per choice alternative at the *aggregate* level.

[26] In a X-fold cross-over validation test, each time a model is estimated based on approximately (100-(100/X))% of all observations to obtain the probabilities for the remaining (100/X)% of the observations; this procedure is repeated X times to obtain classifications for all observations. If X is large (e.g. 10-20), cross-validation is rather time-consuming, but the DTs based on (100-(100/X)) % of all observations resemble the actual DT best. Smaller X's reduce processor time, but the results provide a rougher estimate of the DT's validity. In comparison with validation approaches based on partitioning the data into a fixed training and a fixed validation set, cross-validation has the advantage that the resulting DT is based on all observations rather than a (random) sample of the available data.

For each DT, the evaluation of the combinations of before/after stopping criteria aims to isolate the combination that maximises the percentage of correctly cross-validated observations and the aggregate performance, while at the same time minimising the number of decision rules in the DT. For practical reasons, an all-or-nothing assignment is used for the sensitivity analysis (as opposed to the final DT that uses probabilistic assignment). The results of these analyses for all DTs in the *MERLIN*-system are presented in Appendix 2. For the choice of the travel party (Table A2.1), for instance, the stopping criteria are set at 100 before and 45 after splitting, because (using an all-or-nothing assignment) in the 4-fold cross-validation this DT is able to correctly classify 69.72% of the observations (which almost equals the performance of the models with (many) more rules), while its aggregate performance is even better than that of the model with 80 observations before and 35 after the split of a dicision rule.

8.4 The generated decision table

Given the conceptual and modelling considerations discussed in the previous section, the decision rules for the choice of travel party are induced using the "Exhaustive CHAID"-algorithm available as part of SPSS's AnswerTree® version 2.1 (SPSS, 1998). Since the holidays with "unknown" travel parties are included, all available 7121 observations are used to induce 68 decision rules.

Table 8.1 DT for the choice of travel party

Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Car	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{1}	{1}
Cst	{1}	{2}	{3,4}	{3,4}	{5}	{5}	{5}	{6}	{1,2}	{1,2}
#DT	-	-	(<-,2]	(2,->) + {99}	-	-	-	-	-	-
Age	-	-	-	-	(<-,29]	(29,36]	(36,->)	-	-	-
Lngth	-	-	-	-	-	-	-	-	(<-,1]	(<-,1]
#Hol	-	-	-	-	-	-	-	-	(<-,3]	(3,6]
Alone	.0421	.0000	.1061	.1702	.1209	.1837	.3765	.0645	.0038	.0091
Adult	.8211	.8393	.6364	.4255	.6703	.4898	.5176	.5000	.7681	.7500
SchCh	.0211	.0000	.1515	.0851	.0220	.0408	.0471	.0161	.0190	.0091
OChld	.0000	.0179	.0606	.0213	.0220	.0816	.0118	.2903	.0684	.0273
P9+	.0947	.0357	.0303	.2979	.1099	.1837	.0471	.0968	.1331	.1091
Unkn	.0210	.1071	.0151	.0000	.0549	.0204	.0000	.0323	.0076	.0954
N	95	56	66	47	91	49	85	62	263	220
Rule#	1	2	3	4	5	6	7	8	9	10

The resulting DT is presented in Table 8.1. In this table, the row labelled "Alone" indicates the probability of travelling without travel companions given the conditions presented in the column above. Similarly, "Adult" indicates the probability of selecting adults (20+ years) only, "SchCh" and "OChld" denote the probabilities that schoolgoing children (6-14 years) respectively other children (0-5 or 15+ years) will accompany the traveller, "P9+" indicates the probability that the travel party will consist of at least 9 people (regardless their age), and "Unkn" finally indicates the probability that the travel party will be unknown. Given the size of the DT, the following discussion is restricted to the most important and/or most striking variables that condition the selection process for travel parties. In the discussion, rule numbers (Rule# in the Table 8.1), are indicated by R and the number in question.

Table 8.1 Continued - travel party

Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Car	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Cst	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}
Age	-	(<-,36]	(36,61]	(36,61]	(61,->)	-	-	-	-	-
Lngth	(<-,1]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(2,3]	(2,3]	(2,3]
#Hol	(6,->)	(<-,4]	(<-,4]	(<-,4]	(<-,4]	(4,->)	(4,->)	(<-,3]	(3,->)	(3,->)
Ski	-	-	{0}	{1}	-	-	-	-	-	-
Gndr	-	-	-	-	-	{1}	{2}	-	-	-
Work	-	-	-	-	-	-	-	-	{0,1}	{2,3}
Alone	.0000	.0482	.0000	.0217	.0211	.0526	.0143	.0032	.0811	.0000
Adult	.4444	.7349	.8889	.8696	.7746	.6491	.7571	.9117	.8784	.7831
SchCh	.0417	.0241	.0333	.0217	.0493	.0175	.0143	.0126	.0270	.0120
OChld	.0139	.1446	.0500	.0000	.0211	.0702	.0286	.0536	.0000	.0361
P9+	.1250	.0482	.0278	.0870	.1197	.1754	.0429	.0089	.0135	.1084
Unkn	.3750	.0000	.0000	.0000	.0142	.0352	.1428	.0100	.0000	.0604
N	72	83	180	46	142	57	70	317	74	83
Rule#	11	12	13	14	15	16	17	18	19	20

The most significant condition in the selection of the travel party is the presence of children in the household of the traveller. When there are *no children in the household*, the probability of travelling together with children evidently decreases (R1-35). For holidays by people without children, the possession of a *tourist accommodation with a permanent location* is the next most significant variable. When people own such an accommodation, the likelihood of unknown parties is very high (R26-35) because trip frequencies of these people are often very high (see section 8.3.1). The probability of unknown parties only drops below 50%

Table 8.1 Continued - travel party

Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Peracc	{0}	{0}	{0}	{0}	{0}	{1}	{1}	{1}	{1}	{1}	{1}
Car	{1}	{1}	{1}	{1}	{1}	-	-	-	-	-	-
Cst	{1,2}	{3,4}	{5}	{5}	{6}	-	-	-	-	-	-
Lngth	(3,->)	-	-	-	-	-	(<-,1]	(<-,1]	(<-,1]	(1,2]	(1,2]
#Hol	-	-	(<-,4]	(4,->)	-	-	(<-,6]	(6,->)	(6,->)	(<-,13]	(<-,13]
HHsz	-	-	-	-	-	(<-,1]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]
Sclss	-	-	-	-	-	-	-	(<-,2]	(2,->)	-	-
Prov	-	-	-	-	-	-	-	-	-	{1,5,6, 9,11,12}	{2,3,4, 7,8,10}
Alone	.0145	.1176	.1701	.0833	.0922	.4000	.0000	.0090	.0000	.0444	.2881
Adult	.9217	.6588	.6432	.5833	.5532	.2421	.4889	.0901	.0395	.5333	.0000
SchCh	.0000	.0588	.0207	.0119	.0071	.0105	.0000	.0000	.0000	.0222	.0000
OChld	.0290	.0588	.0332	.0119	.1560	.0211	.0000	.0090	.0000	.0444	.0000
P9+	.0232	.1059	.1286	.1905	.1844	.0526	.1111	.0541	.0000	.1778	.0000
Unkn	.0116	.0001	.0042	.1191	.0071	.2737	.4000	.8378	.9605	.1779	.7119
N	345	85	241	84	141	95	45	111	152	45	59
Rule#	21	22	23	24	25	26	27	28	29	30	31

for tourists who belong to a 2-persons household and when additionally the holiday takes at least 9 days (R33) or the holiday is a medium long holiday, the tourist makes less than 13 holidays per year and the tourist lives in Groningen, Flevoland, Gelderland, South-Holland, North-Brabant or Limburg (R30).

Table 8.1 Continued - travel party

Chld	{1}	{1}	{1}	{1}	{2}	{2}	{2}	{2}	{2}	{2}
Peracc	{1}	{1}	{1}	{1}	{0}	{0}	{0}	{0}	{0}	{0}
Cst	-	-	-	-	{1,2,3,4}	{1,2,3,4}	{1,2,3,4}	{1,2,3,4}	{1,2,3,4}	{1,2,3,4}
Age	-	-	-	-	(<-,42]	(42,->)	-	-	(<-,42]	(<-,42]
Lngth	(1,2]	(2,->)	-	-	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,2]	(1,2]
#Hol	(13,->)	-	-	-	-	-	-	-	-	-
Gndr	-	-	-	-	{1}	{1}	{2}	-	-	-
HHsz	(1,2]	(1,2]	(2,3]	(3,->)	-	-	-	-	-	-
Region	-	-	-	-	{1,3}	{1,3}	{1,3}	{2}	-	-
Fracc	-	-	-	-	-	-	-	-	{0}	{1}
Alone	.0000	.0066	.0430	.0000	.0000	.0000	.0122	.0532	.0390	.0000
Adult	.1053	.5364	.1613	.3889	.1935	.2128	.3049	.2979	.0390	.0920
SchCh	.0000	.0132	.0215	.0185	.6129	.3191	.3049	.1809	.7792	.6092
OChld	.0000	.0066	.0215	.0741	.0000	.2979	.0122	.0319	.0649	.0460
P9+	.0263	.0132	.0215	.0185	.1613	.1489	.3171	.2234	.0649	.2414
Unkn	.8684	.4240	.7312	.5000	.0323	.0213	.0487	.2127	.0130	.0114
N	76	151	93	54	62	47	82	94	77	87
Rule#	32	33	34	35	36	37	38	39	40	41

People from households without children who *do not own a tourist accommodation with a permanent location* often travel with adults only (74.72%; R1-25). Under these conditions, the proportion of this category drops below 50% only when people are divorced or widowed, do not own a car, and make at least 3 day-trips a year (R4), or when the holidays is a short break and the tourists are married or cohabiting and make at least 7 holidays a year (R11).

When there are *very young children (0-5 years) in the household*, the travel party is often unknown when the tourist owns an *accommodation with a permanent location* (R68). When *no such accommodation* is available, the travel party often consists of other than schoolgoing children (74.81%; R56-61), where the proportion of these travel companions only declines below 50% for short breaks by people who are married (R56).

Table 8.1 Continued - travel party

Chld	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Car	-	-	-	-	-	-	-	-	-	-
Cst	{1,2,3,4}	{1,2,3,4}	{1,2,3,4}	{1,2,3,4}	{1,2,3,4}	{5,6}	{5,6}	{5,6}	{5,6}	{5,6}
Age	(42,->)	(<-,42]	(42,->)	(42,->)	-	-	-	(<-,8]	(8,->)	-
Lngth	(1,2]	(2,3]	(2,3]	(2,3]	(3,->)	(<-,1]	(<-,1]	(1,->)	(1,->)	(1,->)
Gndr	-	-	{1}	{2}	-	-	-	-	-	-
HHsz	-	-	-	-	-	-	-	(<-,4]	(<-,4]	(4,->)
Prov	-	-	-	-	-	-	-	{1,2,8, 9,12}	{1,2,8, 9,12}	{1,2,8, 9,12}
Wrkhr	-	-	-	-	-	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
WrkW	-	-	-	-	-	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Csize	-	-	-	-	-	(<-,4]	(4,->)	-	-	-
Alone	.0177	.0208	.0000	.0000	.0113	.0070	.0130	.0000	.0000	.0000
Adult	.2301	.0729	.2553	.0556	.0283	.0140	.0130	.0000	.0000	.0000
SchCh	.4159	.7917	.3830	.7778	.7345	.6084	.5065	.9508	.7938	.6944
OChld	.1681	.0417	.2979	.1667	.1695	.0070	.0130	.0492	.0515	.0093
P9+	.1239	.0625	.0638	.0000	.0452	.3636	.2857	.0000	.1546	.2870
Unkn	.0443	.0104	.0000	.0000	.0112	.0000	.1688	.0000	.0001	.0093
N	113	96	47	54	177	143	77	61	97	108
Rule#	42	43	44	45	46	47	48	49	50	51

If there are *schoolgoing children in the household* of the traveller, again the possession of a tourist accommodation with a permanent location is important. When people own a *tourist accommodation with a permanent location*, the travel party is likely to be unknown for *short breaks* (86.18%; R62-65). For *longer holidays (5+ days*) unknown parties are also likely (40.78%), but under these conditions the party may also include schoolgoing (35.20%) or other (12.29%) children (R66+67).

Table 8.1 Continued - travel party

Chld	{2}	{2}	{2}	{2}	{3}	{3}	{3}	{3}
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Cst	{5,6}	{5,6}	{5,6}	{5,6}	{1}	{2,3,4,5,6}	-	-
Lngth	(1,->)	(1,->)	-	-	(<-,1]	(<-,1]	(1,2]	(1,2]
Prov	{3,7}	{4,5,6,10,11}	-	-	-	-	-	-
Wrkhr	(<-,0]	(<-,0]	(<-,0]	(0,->)	-	-	-	-
WrkW	(<-,0]	(<-,0]	(0,->)	-	-	-	(<-,0] + {9}	(0,1]
Alone	.0000	.0040	.0139	.0385	.0000	.0000	.0000	.0000
Adult	.0000	.0000	.0139	.0897	.1827	.0000	.0153	.0851
SchCh	.8214	.8745	.3056	.1667	.0385	.0556	.0229	.0426
OChld	.0000	.0560	.4514	.3590	.4231	.6574	.8931	.6809
P9+	.0893	.0607	.2014	.2692	.2885	.2037	.0687	.1915
Unkn	.0893	.0048	.0138	.0769	.0672	.0833	.0000	.0000
N	56	247	144	78	104	108	131	47
Rule#	52	53	54	55	56	57	58	59

For people from *households with schoolgoing children who do not own a permanent tourist accommodation*, the civil state is important. People who are *single* or who have *"another" civil state* (mostly children) prefer travel parties with schoolgoing children (64,79%; R47-55). However, when people are not younger than 16 years (R54), or when people work at least 1 hour (R55) parties with other children gain importance. Although to a lesser extent, for people who are *married*, *cohabiting*, *divorced* or *widowed*, parties with schoolgoing children also prevail (55.66%; R36-46), especially for holidays of at least 5 days (65.44%; R40-46).

Table 8.1 Continued - travel party

Chld	{3}	{3}	{2}	{2}	{2}	{2}	{2}	{2}	{3}
Peracc	{0}	{0}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Lngth	(1,2]	(2,->)	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,2]	(2,->)	-
#Hol	-	-	(<-,6]	(6,13]	(13,->)	(13,->)	-	-	-
Gndr	-	-	-	-	{1}	{2}	-	-	-
WrkW	(1,->) \ {9}	-	-	-	-	-	-	-	-
Alone	.0208	.0096	.0000	.0000	.0000	.0000	.0000	.0092	.0274
Adult	.0000	.0287	.0339	.0381	.0000	.0000	.0857	.0367	.0000
SchCh	.0208	.0622	.2373	.0476	.0462	.0000	.3143	.3761	.0137
OChld	.9376	.8373	.0847	.0190	.0154	.0090	.1000	.1376	.4521
P9+	.0208	.0478	.1186	.0095	.0308	.0000	.0857	.0367	.0137
Unkn	.0000	.0144	.5255	.8858	.9076	.9910	.4143	.4037	.4931
N	48	209	59	105	65	111	70	109	73
Rule#	60	61	62	63	64	65	66	67	68

For short breaks by these people, the holiday region of the tourist is important. For tourists from the northern or southern holiday region, parties with schoolgoing children, adults only and/or at least 9 people are preferred by male tourists (R38). For women from these regions, parties with schoolgoing children are preferred by individuals younger than 42 years (R36), while parties with schoolgoing and/or other children and adults only are selected by female tourists of at least 43 years (R37). People from the mid holiday region, finally, prefer parties with adults only, parties of at least 9 people and/or unknown parties (R39).

8.5 Statistics and validation

The previous sections have discussed the induction and the structure of the DT for the choice of travel party. This final section discusses the most important statistics of the DT and several indicators of its predictive abilities. The statistics and validation results for the DT for the choice of travel party are conveniently arranged in Table 8.2.

Table 8.2 Confusion matrix & statistics for the DT *travel party*

	Alone'	Adults'	ChildS'	ChildO'	P9+'	Unkn'	*Total'*
Alone	**.1805**	.4925	.0717	.0654	.1059	.0840	*.0348*
Adults	.0470	**.6977**	.0438	.0546	.0840	.0730	*.3801*
ChildS	.01298	.0865	**.6184**	.1001	.1286	.0535	*.1924*
ChildO	.0174	.1583	.1471	**.5077**	.1080	.0616	*.1310*
P9+	.0366	.3211	.2459	.1408	**.1797**	.0758	*.1007*
Unkn	.0287	.1594	.0638	.0500	.0474	**.6507**	*.1609*
Total	*.0348*	*.3801*	*.1924*	*.1310*	*.1007*	*.1609*	***.5798***

N° of observations	7121	
Stopping criteria	100 before/45 after	
α	0.05	
N° of columns (= rules)	68	
Theta (1 col)	2707 (.3801)[a]	1680 (.2359)[b]
Theta (68 cols)	4999 (.7020)[a]	4129 (.5798)[b]
χ^2_{DT}(df) / Cont. coef.	13138.31 (335) / 0.8053	

[a] Deterministic [b] Probabilistic

To indicate the accuracy of the DT, this table presents two variants of the Theta measure, a chi-square statistic and a contingency coefficient. Basically, the Theta measures indicate the number of correctly classified observations of different model structures and under different assumptions regarding the assignment of observed holidays. Theta(1 col) indicates the number of correctly classified observations of a model that only has one rule (or column in the DT). Basically, this single-rule model (or *null-model*) represents the aggregate distribution of the observations (in this study: the holidays) over the choice alternatives (in this chapter: the travel party-options). In contrast, Theta(#R cols), represents the number of correctly predicted cases when the decision rules that were induced from the data are used, where #R represents the number of decision rules in the DT. As a consequence, the difference between Theta(#R cols) and Theta(1 col) represents the improvement in prediction accuracy of the DT in comparison with the null-model.

Since decision rules can be deterministic as well as probabilistic, two variants of the Theta measures can be calculated for each (single- or #R-column) DT. If a *deterministic* assignment rule is used (this is also called a majority rule[27]), all holidays are assigned to the dominant choice alternative within a column. In the case of the choice of travel party, the dominant choice option at the aggregate level is "adults only" with 2707 observations, comprising 38.01% of all holidays. Hence, the deterministic (or conventional) Theta(1 col) is 2707 (.3801 as a proportion of all observations). Alternatively, the conventional Theta for the full deterministic model can be found in the sensitivity analysis (see Appendix 2). From Table A2.1 it can be learned that 70.20% of the 7121 observations were correctly (re)predicted using the deterministic full-model; hence, the conventional Theta(68 cols), is 4999 (.7020 × 7121). This represents an important improvement (84.67%) compared to the null-model.

MERLIN, however, uses probabilistic DTs. The second variant of theta therefore represents the expected number of correctly predicted cases if a *probabilistic* rule is used to assign observations to the choice alternatives. This measure (T_p) is calculated as:

$$T_p = \sum_{ij} \frac{(f_{ij})^2}{n_j}$$

where f_{ij} is the frequency of the i-th choice alternative in the j-th column and n_j is the total frequency of the j-th column. The resulting probabilistic Theta(1 col) of the DT for the choice of travel party is 1680 (23.59% of the observations), and the probabilistic Theta(68 cols) is 4129 (57.98%), showing an improvement of 145.8% of the predictive ability that can be attributed to the 68 probabilistic decision rules.

27 This variant represents the sum of frequencies of the modal category across the columns of the table and corresponds to the traditional measure in THAID analysis.

The next statistic in Table 8.2 indicating the accuracy of the DT is the well known Chi-squeare (χ^2_{DT}) and the degrees of freedom (df) of this statistic (using probabilistic decision rules). As this statistic is dependent on the sample size, the contingency coefficient, which represents a normalisation of the χ^2_{DT}-statistic, is also given. The contingency coefficient C is given by:

$$C = \sqrt{\frac{\chi_{DT}^{\ 2}}{\chi_{DT}^{\ 2} + N}}$$

where, *N* is the number of observations. This statistic can be interpreted as a measure of the association rate between conditions and response distributions on a zero to one scale because the minimum value of the coefficient is zero (in case of complete statistical independence between the condition and response variables) and the maximum value approximates one. More precisely, in case of complete statistical dependence, the maximum of the contingency value is given by:

$$\max C = \sqrt{\frac{\min((r, \#R) - 1)}{\min((r, \#R) - 1) + 1}}$$

where *r* is the number of rows (in the DTs this is the number of alternatives in the choice set), and, as before, #R represents the number of decision rules (or columns) in the DT. In case of the DT for the choice of travel party, the contingency coefficient is 0.8053 where the maximum for a 6×68 DT equals 0.9258, indicating a rather strong relationship between the conditions of the DT and the response distribution over the travel party options.

Finally, the upper part of Table 8.2 presents the confusion matrix of the probabilistic DT. The elements of this matrix express, for each combination of choice alternatives, the proportion of observations of choice alternative *i* that are classified as choice alternative *i'*. In formula, these elements are expressed as:

$$p_{ii'} = \frac{\sum_j (f_{ij} a_{i'j})}{n_i}$$

where $a_{i'j}$ is the probability of assigning the *i'*-th choice alternative to a case in the *j*-th column, f_{ij} is the actual frequency of the *i*-cases in the *j*-th column, and n_i is the total number of *i*-cases in the sample. The diagonal cells where *i* = *i'*, represent the proportions of correctly classified observations (marked in bold), and the expected proportion of correctly classified cases of the total sample (utmost right-bottom cell of the diagonal) corresponds to the probabilistic Theta(68 cols). Comparing these diagonal cells to the proportion of observations that would be correctly classified without the DT (i.e. with the null-model which would correctly predict the share of the choice alternative at the aggregate level; this share can be found in the column labelled *Total'*) indicates the contribution of the decision-rules to the accuracy of prediction of tourist choices. In the case of the DT for the choice of travel party, for instance, the model correctly predicts 18.05% of the category "alone", as opposed to 3.38% without a model, an improvement of more than 400%. Less spectacularly, but still significantly, the accuracy of prediction for the other travel party categories improve by 78% (parties of 9 or more people) to 304% (unknown travel parties).

8.6 Conclusion and discussion

This chapter was the first in a series of chapters presenting the decision tables (DTs) that describe tourist's profiling decision for holidays. The DT that was discussed in this chapter will be used by the *MERLIN*-system to predict tourists' choices of travel party. As the first chapter in a series, this chapter also formally discussed the CHAID-based algorithm that is used to induce tourist decision rules, and the statistics and confusion matrix that are indicators of the performance of the resulting DT.

The DT for the choice of travel party comprises 68 decision rules that describe the conditions under which the tourist will select one of the six available travel party options (where, due to data limitations, one option is "travel party unknown"). The face validity of the decision rules was satisfactory, with the most decisive conditions being the household composition (i.e. the presence of children in the household) and the possession of a tourist accommodation with a permanent location. When people own such an accommodation, often the travel party is unknown because these people typically have high trip frequencies, while the data collection did not record detailed information on the travel party on every third and subsequent holiday in a particular quarter. With regard to the effect of the household composition on the selection of travel companions, it was found that the presence of children (schoolgoing or otherwise) in the household significantly increases the probability of travelling together with children.

With regard to the validity of the DT, several indicators were discussed. When using probabilistic decision rules, the contribution of the DT to the prediction of travel party choice was significant. Overall, 57.98% of the observed travel party choices are expected to be predicted correctly by the DT against 23.59% by the null-model.

9 Tourist Timing Decisions

This chapter addresses the induction and representation of tourist timing decisions. Three dimensions are identified, including the decision whether or not to leave during a school holiday period, the choice of season and the actual day of departure.

9.1 Introduction

Given the duration and travel party, the next stage in the *MERLIN*-system is the timing of these holidays. The timing of tourist trips is important to the tourism industry in many ways. From a transport point of view, understanding the nature of tourist timing decisions is vital to reducing congestion problems and (fatal) traffic accidents. In France, for instance, "*le changement de juilletists et d'aoûtists*" (i.e., the annual return of July-holidaymakers and the simultaneous departure of August-holidaymakers), is often referred to as "Black Saturday". For this reason, the Netherlands are divided into three so-called school holiday regions that have slightly different school holidays (Stichting Toerisme & Recreatie AVN & Ministry of OC & W, 1997). Tourist timing choices are also relevant to marketing because seasonality may threaten occupancy rates (Bonn *et al*., 1992) or even the viability of the tourism industry of an entire country. The Canadian Tourism Commission, for instance, has undertaken a series of initiatives to position Canada as a four-season tourist destination[28]. Also, trips made in different seasons may be undertaken by tourists with different socio-demographic profiles (Bonn *et al*., 1992) and preference and constraint structures. Understanding these relations is vital to the question of when and how to target different tourist segments.

Evidently, timing decisions have more than one dimension of "peak-" or "off-peak" periods. The next section therefore first identifies these dimensions and the corresponding choice alternatives. Next, following the structure presented in the previous chapter, the choice conditions, the settings of the CHAID-based algorithms, the induced DTs and their validation are discussed.

9.2 The question of when to travel

9.2.1 The choice sets

From a transport and a marketing point of view, at least three timing decisions are relevant to the tourism industry. Firstly, seasonality is a major challenge for

[28] Source: http://canadatourism.com/archive/tourism/success/story/html

accommodations and destinations, where usually summer or winter (in case of winter sports areas or destinations on the Southern Hemisphere) are denoted as the high season. Bonn *et al.* (1992), for instance, identified differences between fall-, winter-, spring- and summer-visitors to a popular coastal resort destination area that can be applied to marketing strategies.

Secondly, school holiday periods are relevant to both occupancy rates and congestion. In addition, given the choice of travel party or restrictions imposed on free days, travelling during school holiday periods may be inevitable. However, this often comes at the cost of high season prices, and over-crowded roads, destinations and accommodations.

Third, and finally, the actual date of departure is important to congestion problems because many holidaymakers leave the day or the day after the weekend commences, i.e. on Friday or Saturday. On the other hand, people may leave on off-peak days to avoid congestion and high fares on, for instance, ferries and toll roads. Understanding the characteristics of the people who depart on Fridays or Saturdays and/or the holidays that start on these peak days may help policy makers to develop plans to use the available transport system more effectively.

Based on these demand- and supply-side considerations, the MERCIN-system comprises three timing decisions, including "during or outside school holiday periods", the season, and the departure date. The decision variable "*during or outside school holiday periods*" has two choice alternatives, including (1) during a school holiday period, and (0) outside a school holiday period. Based on the prescribed summer holidays for schools, free days/extended weekends as a result of national and religious holidays, and the recommended dates for spring, autumn and Christmas holidays in the Netherlands in the period between December 1997 and November 1998 (the 1998 "CVO-year"), the first choice alternative is defined as (mainly based on Stichting Toerisme & Recreatie AVN & Ministry of OV & W, 1997):

- Christmas holidays: Dec. 19th, 1997 (Fri) through Jan. 3rd, 1998 (Sat);
- Spring holidays: Feb. 2nd, 1998 (Fri) through Feb. 28th, 1998 (Sat);
- Easter Weekend: Apr. 9th, 1998 (Thu) through Apr. 13th, 1998 (Tue);
- May Holiday: Apr. 24th, 1998 (Fri) through May 5th, 1998 (Tue);
- Ascension Day: May 20th, 1998 (Wed) through May 23rd, 1998 (Sat);
- Whit Weekend: May 29th, 1998 (Fri) through June 1rst, 1998 (Mon);
- Summer holidays: July 3rd, 1998 (Fri) through Sept. 5th, 1998 (Sat); and
- Autumn holidays: Oct. 16th, 1998 (Fri) through Oct. 31 rst, 1998 (Sat).

Tourist trips are classified based on the date of departure. Holidays starting outside the above mentioned dates are classified as *"outside school holiday period"*. For example, if a tourist left for a 4-day holiday on Friday September 4, this is classified as a short break in a school holiday period, whereas if the departure date had been Sunday September 6, this had been a short break outside school holiday period.

Table 9.1 Distribution of the tourist trips with regard to season

winter	spring	summer	autumn
846 (11.9%)	2551 (35.8%)	2612 (36.7%)	1112 (15.6%)

The second dimension of tourist timing decisions is of the choice of *season*. Obviously, the choice of season for each tourist trip includes the choice alternatives winter, spring, summer and autumn. Based on the schedule of school holidays in 1997/1998 and the available data, *winter* is defined as the period from December 1st, 1997 through February 28, 1998 inclusive. *Spring* is set from March 1st through to July 2nd, 1998. *Summer* is defined as the 9-week period during which (primary and secondary) schools in the Netherlands may be closed for the summer holiday starting at July 3rd and ending at September 5th, 1998. *Autumn*, finally, runs from September 6th through November 30th, 1998 inclusive.

Using the above definitions, the distribution of the 7121 trips across these actions is displayed in Table 9.1. Although the summer season extends over a 9-week period only, it is still the dominant season for holidaymaking. Unlike the almost equally popular spring season, however, the summer high season is not further divided by, for instance, school holidays. It is decided, therefore, to further divide the summer season into *three equal periods* of three weeks. In the first tree weeks (*summer-begin*: July 3rd-23rd), schools throughout the country subsequently close their doors for the summer holidays. In the second three weeks (*summer-mid*: July 24th through Aug. 13th) all schools in the Netherlands are closed, and the final tree weeks (*summer-end*: Aug. 14th through Sept. 5th), region by region schools open their doors again. This second choice of season is modelled conditional upon the first choice (of the main season). In conclusion, the choice of season is described by a set of two hierarchical DTs, and *MERLIN* considers the following categories for the choice of season (including coding): (1) winter; (2) spring; (3) summer-begin; (4) summer-mid; (5) summer-end; and (6) autumn.

The third and final timing dimension for tourist trips is the *date of departure* in terms of the day of the week. As mentioned earlier, this choice facet is important with regard to congestion control because many people (56.9% in 1998) depart on the day or the day after the weekend commences, i.e. on Friday or Saturday. For this decision, therefore, *MERLIN* considers two choice options (including coding): (1) departure on Friday of Saturday; and (0) departure on another day.

Obviously, the order in which these timing decisions are considered will affect the decision rules that can be identified because previous decisions can be included as conditions in subsequent decisions. Presumably, the day of departure is often decided only when all other facets of the travel decision have been contemplated, including the choice of destination, accommodation and the mode of transport. A hotel or aeroplane reservation, for instance, will leave no or little freedom for the traveller to move his or her date of departure. A domestic car–based camping holiday, on the other hand, can be delayed or moved up more easily. As argued in chapter 2, therefore, the *MERCIN*-system considers the date of departure conditional upon all other profiling decisions (except expenditures).

With regard to the profiling order of the choices of season and the decision whether or not to leave during a school holiday period, the relationship between of the decisions is assumed to be bi-directional, where, in some cases the season may be considered first, whereas in other cased the “school holiday or not” dimension may be contemplated first. Also, the “season” and “school holiday period or not” choices may be taken simultaneously. Unfortunately, based on empirical timing data it is not possible to determine which decision is taken first. Alternatively, therefore, a more empirical approach is adopted, and a small simulation experiment was conducted to answer to question: Which timing decision should be modelled first, “season” or “during school holiday period or not”, in order to obtain the set of decision rules that best (re-)predicts tourist timing choices? Hence, two sets of two probabilistic DT’s were induced, the first of which described the choice of season (where summer was considered as one compound category) conditional upon the choice regarding school holiday periods. In contrast, the second set of DT’s described the “school holiday period or not”-choice conditional upon the choice of season. Using random numbers, the average number of correctly predicted observations (combined as well as on separate dimensions) and the performances at the aggregate level were compared after 15 runs. The average numbers of correctly predicted observations did not differ much. Eventually, the first set of DTs was selected because it performed slightly better at the aggregate level. Thus, the *MERCIN*-system considers the choice of “school holiday period or not” first, and the choice of season is modelled conditional upon this choice.

9.2.2 The condition variables

The previous chapter has introduced several condition variables that are common to all choice facets (see also Table A1.1 in Appendix 1). Since the choice of travel party precedes the tourists’ timing choices, this variable is also entered into the

CHAID-based algorithm as a condition for tourist timing decisions (Table A1.2). In addition, several condition variables related to the *profiling order of the tourist trips within a tourist's trip program* are included. In chapter 5 it has been argued that the order in which an individual profiles his or her tourist trips affects the outcome of the profiling process because decisions on other trips may affect decisions on trips yet to come. Also, based on conceptual and practical considerations, the so-called *profiling priority* of a trip within an individual's trip program is determined first by the duration of a trip (the longer the trip, the higher the priority) and second, when the duration is identical, on the travel party. With regard to the latter criterion, trips *with schoolchildren* have the highest profiling priority because these travelling companions have fixed (school and holiday) agenda's; trips for *single persons*, on the other hand, have lowest-but-one priority because the traveller does not have to anticipate other people's schedules (when the travel party is *unknown*, the trip has the lowest priority). For similar reasons, trips with *children between 0-5 and 15-19 years* old, travel parties with *adults only* (> 20 year) and with travel parties of *9 or more people*, respectively, sit in-between.

Three condition variables derived from the profiling priority of a holiday are entered into the CHAID-based for tourist timing choices (and all subsequent choice facets; see Table A1.2). The first condition variable distinguishes between (1) the most important holiday within the profiling process; (2) the first but one most important trip; (3) the first but two most important trip; and (4) trips that are not part of the three most important trips. The second related condition variable is the number of trips that still remain to be profiled (including the trip under consideration). This variable is the opposite of the profiling priority, and is calculated as the total number of trips in an individual's tourist trip program (#Hol) minus the profiling priority of the trip under consideration plus one. The third and final condition variable is the total number of days of trips with higher profiling priorities that have already been profiled.

In addition to these condition variables common to all tourist timing (and subsequent) decisions, each timing decision has as set of conditions that is specific to that choice. First, for the choice of season, the condition "during school holiday period or not" is included because this choice has been taken previously. Similarly, for the choice of departure date, the other timing choices as well as the choices of destination, accommodation and transport mode are included as condition variables (Table A1.2). Secondly, for each timing dimension, several *summary condition variables* are included. These variables summarise the decisions on the choice facet under considerations for holidays with higher profiling priorities both in terms of the number of trips that are associated with each choice alternative, as well as the number of days that are involved. For the first timing dimension, there are four summary condition variables (Table A1.3), including the number of days of more important trips already scheduled in school holiday periods (#Dih), the number of days of more important trips already scheduled outside school holiday periods (#Doh), the number of more important trips already scheduled in school holiday periods (#Tih), and the number of more important trips already scheduled outside school holiday periods (#Toh). By definition, for tourist trips with the highest profiling priority within an individual's tourist trip program, these summary values are always zero. If, for instance, the tourist trip with the highest profiling priority is a long holiday, and this holiday is scheduled during a school holiday period, the values of the summary values for the first but one most important holiday are: #Dih = 22; #Doh = 0; #Aih = 1; #Aoh = 0. Similarly, for the choice of season there are eight summary condition variables representing the number (of days of) trips already scheduled in each of the four seasons (Table A1.4). The numbers of trips already scheduled during the four seasons are also entered into the CHAID-based algorithm for the choice of the summer season period. In addition, three summary variables represent the number of trips already scheduled in the three parts of the summer holiday season (Table A1.5). The amounts of days that are associated with these decisions are not included in order to limit the number of condition variables. Finally, for the choice of the departure date, four summary condition variables are included (Table A1.6).

9.2.3 The stopping criteria of the CHAID-based algorithm

Finally, given the choice sets and the condition variables for the timing choice dimensions, the stopping criteria for the CHAID-based algorithm have to be specified. For each DT, the α-level for predictor eligibility is set at 5%. Furthermore, for each DT several combinations of before/after stopping criteria were examined (see Appendix 2). Based on these analysis, the stopping criteria for the choice of "school holiday period or not", are set at 80 before and 35 after splitting, because (using an all-or-nothing assignment) this setting performs best at the 4-fold cross-validation while the aggregate performance equals or even exceeds that the other settings (Table A2.2). Similarly, for the choice of season, the 60 before/25 after setting is selected because less strict criteria come at the cost of a decreasing aggregate performance (Table A2.3). The algorithm for the choice of summer season period (Table A2.4) is set to have at least 80 observations before and 35 observations after splitting. This model has one decision rule more than the 100 before/45 after criteria, but it performs slightly better on the cross-validation test. Finally, for the choice of departure date (Table A2.5) the 100 before/45 after criteria are selected because this model outperforms all other models while using a minimum number of rules.

9.3 The generated decision tables

This section discusses the DTs for tourist timing choices. The structures of the DTs generated by the CHAID-based algorithm are presented in the Tables 9.2 through 9.5. The next sections discuss the tables for these timing decision. Given the size of some of the DTs, the discussion is restricted to the most important and/or most striking variables that condition timing decisions. As before, rule numbers (Rule# in the DTs), are indicated by R and the number in question.

9.3.1 The DT for the choice of "school holiday period or not"

The generated DT for the choice of "during school holiday period or not" comprises 46 decision rules and is shown in Table 9.2. In this table "Yes" indicates that a holiday will be scheduled during a school holiday period (with a certain probability), whereas "No" indicates the selection of a period outside school holidays. As expected, the presence of children in the household is very important in the decision to travel during a school holiday period or not. If there are *schoolchildren (6-17 years) in the household* (R29-46), the majority of holidays are scheduled during school holiday periods. Exceptions occur, however, in case of short breaks by large parties with profiling priority 2 of 3 (R40), when the travel party consists of adults only, or when the tourist travels alone (R36), or when people have already allocated many days during school holiday periods (R45+46).

Table 9.2 DT for the choice of "during school holiday period or not"

Chld	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}
Prio4	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]
Party	{1,2,5}	{1,2,5}	{1,2,5}	{1,2,5}	{1,2,5}	{1,2,5}	{1,2,5}	{3,99}	{4}	{4}
Age	(<-,29]	(<-,29]	(29,53]	(29,53]	(53,->)	(53,->)	(53,->)	-	-	-
Lngth	(<-,2]	(2,->)	-	-	-	-	-	-	(<-,2]	(2,->)
Wrkrs	-	-	{0}	{1}	-	-	-	-	-	-
Fracc	-	-	-	-	{0}	{1}	{1}	-	-	-
Csize	-	-	-	-	-	(<-,4]	(4,->)	-	-	-
No	.4844	.2179	.4796	.2644	.5808	.5287	.2692	.1078	.4658	.2679
Yes	.5156	.7821	.5204	.7356	.4192	.4713	.7308	.8922	.5342	.7321
N	64	156	442	87	396	87	52	102	161	224
Rule#	1	2	3	4	5	6	7	8	9	10

Table 9.2 Continued -"during school holiday period or not"

Chld	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}
Prio4	(1,3]	(1,3]	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(1,->)	(1,->)	(1,->)	(1,->)
Party	{1,3}	{99}	{1,3,99}	{1,3,99}	{1,3,99}	{1,3,99}	{1,3,99}	{2,4,5}	{2,4,5}	{2,4,5}	{2,4,5}
Wrkrs	-	-	-	-	-	-	-	{0}	{0}	{0}	{0}
Fracc	-	-	-	-	-	-	-	{0}	{0}	{0}	{0}
#Dih	-	-	(<-,40]	(<-,40]	(<-,40]	(40,->)	(40,->)	-	-	-	-
Perac	-	-	{0}	{0}	{1}	-	-	-	-	-	-
#Dnh	-	-	(<-,18]	(18,->)	-	-	-	-	-	-	-
#Dtot	-	-	-	-	-	(<-,54]	(54,->)	-	-	-	-
Ski	-	-	-	-	-	-	-	{0}	{0}	{1}	{1}
Car	-	-	-	-	-	-	-	{0}	{1}	-	-
#Tnh	-	-	-	-	-	-	-	-	-	(<-,1]	(1,->)
No	.4359	.2564	.2462	.5318	.4870	.7381	.5628	.5113	.6305	.5873	.3478
Yes	.5641	.7436	.7538	.4682	.5130	.2619	.4372	.4887	.3695	.4127	.6522
N	117	117	65	47	269	42	199	133	682	126	46
Rule#	11	12	13	14	15	16	17	18	19	20	21

For *households without schoolgoing children* (R1-28), the profiling priority is the most significant conditioning factor because the most important holiday is scheduled during school holiday periods more often (R1-10).

For households without schoolchildren, a narrow majority of holidays with lower profiling priorities are not scheduled during school holiday periods (52.6%; R11-28). Under these conditions, the *travel party* constitutes the next significant condition variable. Basically, parties of single travellers, with schoolchildren and "unknown" travel parties prefer travelling during school holidays (53.7%; R11-17) whereas parties with adults only, with other children (0-5 or 15+ years) or with 9 or more members prefer to avoid these periods (55.3%; R18-28).

Table 9.2 Continued -"during school holiday period or not"

Chld	{1}	{3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{2}	{2}	{2}
Prio4	(1,->)	(1,->)	(1,->)	(1,->)	(1,->)	(1,->)	(1,->)	(<-,1]	(<-,1]	(<-,1]
Party	{2,4,5}	{2,4,5}	{2,4,5}	{2,4,5}	{2,4,5}	{2,4,5}	{2,4,5}	-	{1,3,99}	{2,4,5}
Lngth	-	-	-	-	-	-	-	(<-,1]	(1,3]	(1,3]
Wrkrs	{0}	{0}	{0}	{0}	{0}	{1}	{1}	-	-	-
Fracc	{1}	{1}	{1}	{1}	{1}	-	-	-	-	-
#Dih	(<-,11]	(<-,11]	(11,17]	(17,24]	(24,->)	-	-	-	-	-
SClss	-	-	-	-	-	(<-,1]	(1,->)	-	-	-
No	.5068	.6618	.6084	.3934	.5577	.2785	.5192	.4697	.0273	.1448
Yes	.4932	.3382	.3916	.6066	.4423	.7215	.4808	.5303	.9727	.8552
N	294	68	166	122	156	79	156	66	439	145
Rule#	22	23	24	25	26	27	28	29	30	31

Table 9.2 Continued -“during school holiday period or not”

Chld	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Prio4	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,3]	(1,3]	(1,3]	(1,3]	(1,3]	(1,3]
Party	{1,3,4,99}	{1,3,4,99}	{2,5}	-	{1,2}	{3,4,99}	{3,4,99}	{3,4,99}	{5}	{5}
Lngth	(3,4]	(3,4]	(3,4]	(4,->)	-	-	-	-	(<-,1]	(1,->)
Fracc	{0}	{1}	-	-	-	-	-	-	-	-
Prov	-	-	-	-	-	{1,2,3,12}	{4,7,8,9}	{5,6,10,11}	-	-
#Hol	-	-	-	-	-	-	-	-	-	-
No	.0594	.0078	.1429	.0789	.5313	.4235	.1672	.2462	.5140	.2688
Yes	.9406	.9922	.8571	.9211	.4687	.5765	.8328	.7538	.4860	.7312
N	101	256	35	38	96	85	323	195	107	93
Rule#	32	33	34	35	36	37	38	39	40	41

However, within the first group of travel parties, people prefer travelling outside school holidays more frequently for less important holidays (“not in top-3”) when more than 40 days have been scheduled during school holidays (R16+17), or when 40 days or less have been scheduled during school holidays, more than 18 days have been scheduled outside these periods and the tourist does not own a permanent tourist accommodation (R14). Exceptions to the preference for avoiding school holiday period of the second group of travel parties, finally, include people who are not restricted in going on holiday due to prescribed holiday regulations at work, who do not own a tourist accommodation without a permanent place, and who have already scheduled at least two holidays outside school holiday periods (R21). In addition, people who are not restricted by prescribed holidays at work, who own a non-permanent tourist accommodation and who have already scheduled 18 to 24 days during school holiday periods, also prefer school holiday periods more frequently (R25). The final exception includes people from the highest social class who are restricted in going on holiday due to prescribed holiday regulations at work (R27).

Table 9.2 Continued -“during school holiday period or not”

Chld	{2}	{2}	{2}	{2}	{2}
Prio4	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)
Party	{1,2,3,5}	{4,99}	-	-	-
Fracc		-	-	{0}	{1}
#Dih		-	(<-,40]	(40,->)	(40,->)
#Hol	(<-,6]	(<-,6]	(6,->)	(6,->)	(6,->)
No	.4595	.2000	.4141	.5823	.7344
Yes	.5405	.8000	.5859	.4177	.2656
N	111	55	99	158	64
Rule#	42	43	44	45	46

9.3.2 The DT for the choice of season

The DT for the choice of season comprises a set of 88 exhaustive and exclusive decision rules that are shown in Table 9.3.

Table 9.3 DT for the choice of season

HolP	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Party	{1}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Ski	-	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
#Tspr	-	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
#Daut	-	(<-,10]	(<-,10]	(<-,10]	(<-,10]	(<-,10]	(<-,10]	(<-,10]	(<-,10]
Lngth	-	(<-,1]	(1,4]	(1,4]	(1,4]	(1,4]	(1,4]	(1,4]	(4,->)
#Twnt	-	-	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	-	-
#Dsum	-	-	(<-,12]	(<-,12]	(<-,12]	(<-,12]	(<-,12]	(12,->)	-
Age	-	-	(<-,42]	(42,53]	(53,61]	(61,->)	-	-	-
Winter	.1842	.2214	.1293	.0556	.0676	.0857	.0667	.2211	.3611
Spring	.4474	.4656	.4741	.6619	.7973	.6214	.2667	.5895	.3889
Summer	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Autumn	.3684	.3130	.3966	.2825	.1351	.2929	.6666	.1894	.2500
N	114	131	116	108	74	140	30	95	36
Rule#	1	2	3	4	5	6	7	8	9

Table 9.3 shows that the previous timing decision is the most significant predictor for the choice of season. This is explained best by the fact that holidays that are scheduled outside school holiday periods can never be pursued during the summer season (R1-38).

Table 9.3 Continued - season

HolP	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Party	{2}	{2}	{2}	{2}	{2}	{2}	{3}	{3}	{3}
Ski	{1}	-	{0}	{1}	-	-	{0}	{0}	{0}
#Tspr	(<-,0]	(<-,0]	(0,->)	(0,->)	(0,->)	(0,->)	-	-	-
#Daut	(<-,10]	(10,->)	(<-,0]	(<-,0]	(0,->)	(0,->)	-	-	-
Lngth	-	-	-	-	-	-	(<-,1]	(<-,1]	(1,->)
Prio4	-	-	-	-	(<-,3]	(3,->)	-	-	-
Fracc	-	-	-	-	-	-	{0}	{1}	-
Winter	.2615	.3000	.1250	.3019	.3415	.1474	.4000	.1522	.1143
Spring	.5154	.6571	.3984	.2264	.3902	.5684	.3714	.6739	.8286
Summer	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Autumn	.2231	.0429	.4766	.4717	.2683	.2842	.2286	.1739	0.0571
N	130	70	256	53	41	95	35	46	70
Rule#	10	11	12	13	14	15	16	17	18

Given the choice to travel *outside school holidays*, spring is the most popular season, and the travel party is the most significant condition to explain the choice between the three remaining seasons. *Single travellers* variably plan during the winter (18.42%), spring (44.74%) and autumn (36.84%) (R1). For *parties of adults only* (R2-15), on the other hand, several other condition variables determine the choice of season. Of these condition variables, the number of holidays already scheduled during the spring is the most important. When no holidays have been planned for spring yet, this season is by far the most preferred option (R2-11), except for extra long holidays that are scheduled more evenly throughout the year (R9). In addition, under these conditions, autumn is the most preferred season when the holiday takes between 5 and 28 days, the tourist does not own skis, at least one day has been planned during the winter, less than 11 days have been planned for autumn and less than 13 days have been planned for the summer yet (R7). When one or more holidays have already been scheduled for the spring, the other two seasons gain importance (R12-14), except for holidays with a low profiling priority (R15). These rules indicate a preference for spreading holidays throughout the year.

Table 9.3 Continued - season

HolP	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Party	{3}	{4,5}	{4,5}	{4,5}	{4,5}	{4,5}	{4,5}	{4,5}	{4,5}
Ski	{1}	-	-	-	-	-	-	-	-
Lngth	-	-	-	(<-,1]	(1,->)	(1,->)	-	-	-
HHsz	-	(<-,2]	(2,3]	(<-,3]	(<-,3]	(<-,3]	(<-,3]	(3,->)	(3,->)
Prio4	-	(<-,1]	(<-,1]	(1,3]	(1,3]	(1,3]	(3,->)	-	-
#Tzmr	-	-	-	-	(<-,0]	(0,->)	-	-	-
Fracc	-	-	-	-	-	-	-	{0}	{1}
#Dspr	-	-	-	-	-	-	-	(<-,0]	(<-,0]
Winter	.2800	.2034	.0882	.2115	.3256	.4545	.1042	.1869	.1059
Spring	.4000	.3898	.6618	.4231	.3256	.4182	.4167	.6542	.8353
Summer	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Autumn	.3200	.4068	.2500	.3654	.3488	.1273	.4791	.1589	.0588
N	25	59	68	104	43	55	48	107	98
Rule#	19	20	21	22	23	24	25	26	27

For *holidays outside school holidays by travel parties with schoolchildren* (R16-19), skis, the trip duration and the possession of a non-permanent accommodation determine the preferences for the three season. Naturally, the possession of skis increases the preference for the winter season (R19). However, people without skis also have a strong preference for this season when it comes to short breaks (R16).

Table 9.3 Continued - season

HolP	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Party	{4,5}	{4,5}	{99}	{99}	{99}	{99}	{99}	{99}	{99}
HHsz	(3,->)	(3,->)	-	-	-	-	-	-	-
#Twnt	(<-,0]	(0,->)	-	-	-	-	-	-	-
#Taut	-	-	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]
#Dwnt	-	-	(<-,10]	(<-,10]	(<-,10]	(<-,10]	(<-,10]	(<-,10]	(10,->)
#Tzmr	-	-	(<-,1]	(<-,1]	(<-,1]	(1,->)	(1,->)	(1,->)	-
#Dspr	(0,->)	-	-	-	-	(<-,17]	(17,->)	-	-
SClss	-	-	(<-,1]	(1,->)	(1,->)	-	-	-	-
#Tp	-	-	-	-	-	(<-,5]	(<-,5]	(5,->)	-
Prov	-	-	-	{1,2,4,5,7,8, 9,10,11,12}	{3,6}	-	-	-	-
Winter	.1566	.0714	.0263	.0000	.0000	.0000	.0000	.0303	.2745
Spring	.5542	.5000	.7895	.9933	.8000	.7755	.3429	.8485	.6275
Summer	.0000	.0000	.0000	.0000	.2000	.0000	.0000	.0000	.0000
Autumn	.2892	.4286	.1842	.0067	.0000	.2245	.6571	.1212	.0980
N	83	28	38	150	25	49	35	66	51
Rule#	28	29	30	31	32	33	34	35	36

Outside the school holiday periods, the choice of season of *travel parties with other children* (0-5 or 15+ years) and *travel parties of at least 9 people* (R20-29) are conditioned by the household size. More specifically, larger households (4 or more member) have a stronger preference for the spring season (R26-29) compared to smaller households.

Table 9.3 Continued - season

HolP	{0}	{0}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Party	{99}	{99}	-	-	-	-	-	-	-
Ski	-	-	-	-	-	-	{0}	{1}	-
#Daut	-	-	-	-	-	-	(<-,10]	(<-,10]	(10,->)
Lngth	-	-	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,2]	(1,2]	(1,2]
#Taut	(1,->)	(1,->)	-	-	-	-	-	-	-
Urban	(<-,2]	(2,->)	-	-	(<-,3]	(3,->)	-	-	-
#Tzmr	-	-	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Chld	-	-	{1}	{1}	{1}	{1}	{1}	{1}	{1}
#Dtot	-	-	(<-,12]	(12,18]	(18,->)	(18,->)	-	-	-
Winter	.1053	.1489	.0776	.2963	.0000	.1154	.0891	.1842	.2381
Spring	.1316	.6382	.2672	.0376	.2800	.4231	.1336	.1579	.2619
Summer	.0000	.0000	.5259	.5556	.6600	.2692	.7126	.5000	.3810
Autumn	.7631	.2129	.1293	.1105	.0600	.1923	.0647	.1579	.1190
N	29	47	116	27	50	26	247	38	42
Rule#	37	38	39	40	41	42	43	44	45

Unknown travel parties (R30-38), finally, predominantly prefer the spring season. Two exceptions stand out. First, people from (strongly) urbanised areas who have already scheduled at least one holiday during the autumn, have a strong preference for the latter season (R37). Second, people who have already scheduled at least 18 days during the spring season, at least one holiday during the summer, and who only have 5 or less holidays left to consider also select autumn more often (R34). Note also that under these conditions, the winter season is avoided altogether. This is probably due to the fact that these holidays are often made by people with high holiday frequencies[29] and who apparently aim to spread these trips across the spring, the summer and the autumn season.

Table 9.3 Continued - season

HolP	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Ski	{0}	{0}	{1}	{1}	-	-	-	-	-
Lngth	(2,4]	(2,4]	(2,3]	(3,4]	(2,4]	(4,->)	(4,->)	(<-,1]	(<-,1]
#Taut	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	-	-	-	-
#Tzmr	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{2}	{3}
Fracc	-	-	-	-	-	-	-	{0}	{0}
Work	-	-	-	-	-	{0,1,2}	{3}	-	-
Csize	(<-,6]	(6,->)	-	-	-	-	-	-	-
Winter	.0309	.0556	.2340	.0600	.1481	.0000	.0769	.1795	.0357
Spring	.0949	.1111	.0638	.1400	.3704	.0606	.4359	.0769	.2143
Summer	.8653	.7407	.6383	.7800	.4444	.8182	.4615	.6410	.5000
Autumn	.0089	.0926	.0639	.0200	.0371	.1212	.0257	.1026	.2500
N	453	54	47	50	27	33	39	39	28
Rule#	46	47	48	49	50	51	52	53	54

Evidently, for the holidays that are scheduled *during a school holiday period* (R39-88), summer is the main season. The most significant condition explaining the choice of season in this segment is the number of holidays already scheduled during the summer. When *no holidays have been planned for the summer yet*, summer is by far the most popular alternative (R39-61). Under these conditions, the presence of children in the household is the most important condition to further influence the choice of season. For households without children (R39-52), summer is still the most important season, but less dominantly than for households with children (R53-61).

[29] Information on the travel party is not always available because the CVO-quarterly measurements do not record the composition of the travel party for the third (and subsequent) holiday(s) in each quarter; see chapter 8.

Table 9.3 Continued - season

HolP	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Party	-	-	-	-	-	-	-	-	{1,99}
Ski	-	-	{0}	{1}	-	-	-	-	-
Lngth	(<-,1]	(1,2]	(2,3]	(2,3]	(3,->)	(3,->)	(3,->)	-	-
Prio4	-	-	-	-	-	-	-	(<-,3]	(<-,3]
#Tzmr	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)
Chld	{2,3}	{2,3}	{2,3}	{2,3}	{2}	{2}	{3}	-	-
#Dsum	-	-	-	-	-	-	-	(<-,12]	(12,24]
Fracc	{1}	-	-	-	{0}	{1}	-	-	-
Winter	.0217	.0503	.0177	.1765	.0000	.0000	.0282	.1699	.0364
Spring	.4348	.1038	.0227	.0588	.0238	.0000	.0000	.3127	.4727
Summer	.5217	.7358	.9419	.7647	.9762	1.000	.9577	.4015	.4364
Autumn	.0218	.1101	.0177	.0000	.0000	.0000	.0141	.1159	.0545
N	46	318	396	34	126	291	71	259	55
Rule#	55	56	57	58	59	60	61	62	63

For households without children, the duration of the holiday, the number of (days of) holidays during autumn, the working situation, and the number of inhabitants of the city of residence condition the preferences for the seasons. For households with children, finally, summer is the most preferred season (> 70%), except for short breaks (R53-55).

Table 9.3 Continued - season

HolP	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Party	{2,3, 4,5}	{2,3, 4,5}	{2,3, 4,5}	{2,3, 4,5}	{2,3, 4,5}	{2,3, 4,5}	{2,3 ,4,5}	-
Ski	{0}	{0}	{0}	{0}	{1}	{1}	-	{0}
#Tspr	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	-
Lngth	(<-,2]	(<-,2]	(<-,2]	(2,->)	-	-	-	-
Prio4	(<-,3]	(<-,3]	(<-,3]	(<-,3]	(<-,3]	(<-,3]	(<-,3]	(<-,3]
#Tzmr	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)
#Dsum	(12,24]	(12,24]	(12,24]	(12,24]	(12,24]	(12,24]	(12,24]	(24,->)
SClss	(<-,2]	(2,->)	(2,->)	-	-	-	-	-
Gndr	-	-	-	-	{1}	{2}	-	-
Region	-	{1,2}	{3}	-	-	-	-	-
Winter	.2201	.1299	.3556	.2157	.2188	.5000	.1616	.1869
Spring	.4498	.7403	.4000	.4314	.4688	.1389	.3535	.6262
Summer	.1531	.0390	.0444	.2941	.1875	.2778	.2525	.1215
Autumn	.1770	.0908	.2000	.0588	.1249	.0833	.2324	.0654
N	209	77	45	51	32	36	99	107
Rule#	64	65	66	67	68	69	70	71

Table 9.3 Continued - season

HolP	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Party	-	{1,4}	{2}	{2}	{2}	{3}	{5}	{99}	{99}
Ski	{1}	-	-	-	-	-	-	-	-
#Tspr	-	-	(<-,1]	(1,->)	-	-	-	(<-,1]	(<-,1]
#Daut	-	-	(<-,0]	(<-,0]	(0,->)	-	-	(<-,10]	(<-,10]
Prio4	(<-,3]	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)
#Tzmr	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)
#Dsum	(24,->)	-	-	-	-	-	-	-	-
Peracc	-	-	-	-	-	-	-	{0}	{1}
Winter	.4286	.1373	.0870	.2500	.1633	.2619	.0930	.0400	.1667
Spring	.4286	.3333	.5435	.1786	.4082	.2381	.4651	.0400	.5417
Summer	.0286	.4118	.1522	.1429	.3673	.2381	.3256	.9200	.2292
Autumn	.1142	.1176	.2173	.4285	.0612	.2619	.1163	.0000	.0624
N	35	51	47	28	49	42	43	25	48
Rule#	72	73	74	75	76	77	78	79	80

When *at least one holiday has already been scheduled during the summer season*, the probability of scheduling another holiday during this season decreases dramatically (R62-88). In this case, the priority of the holiday within the trip program is important.

Table 9.3 Continued - season

HolP	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Party	{99}	{99}	{99}	{99}	{99}	{99}	{99}	{99}
#Tspr	(1,5]	(1,5]	(1,5]	(1,5]	(1,5]	(5,->)	(5,->)	-
#Daut	(<-,10]	(<-,10]	(<-,10]	(<-,10]	(<-,10]	(<-,10]	(<-,10]	(10,->)
Prio4	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)
#Tzmr	(0,->)	(0,1]	(1,->)	(0,1]	(1,->)	(0,1]	(1,->)	(0,->)
Work	{0}	{1,2}	{1,2}	{3}	{3}	-	-	-
Winter	.0000	.0000	.0000	.0667	.0816	.0000	.0000	.0294
Spring	.8500	.5143	.2632	.5000	.2245	.4483	.0575	.2353
Summer	.1250	.4571	.7368	.4333	.6939	.5517	.9195	.4118
Autumn	.0250	.0286	.0000	.0000	.0000	.0000	.0230	.3235
N	40	35	38	30	49	29	87	34
Rule#	81	82	83	84	85	86	87	88

If the holiday is one of the three most important holidays within a tourist's trip program (R62-72), the probability of scheduling another holiday during this season decreases even further. Under these conditions, the number of days already scheduled during the summer, the travel party, the number of days already scheduled in the spring, the possession of skis, the social class, gender and the holiday region further structure the choice of season. If, on the other hand, the holiday is not one of the three most important holidays within the tourist's trip program (R73-88), the travel party is the next most significant condition variable, where unknown travel parties are more likely to select the summer season (R79-88).

9.3.3 The DT for the choice of summer season period

The DT for the choice of one of the three parts of the summer season is induced using the 2612 observed holidays in the CVO-data set that were scheduled during the summer. It comprises 24 decision rules and is presented in Table 9.4.

Table 9.4 DT for the choice of summer season period

Region	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}
#Tspr	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Chld	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{1,3}	{2}	{2}
Lngth	{1}	{2,3}	{2,3}	{2,3}	{4,5}	{4,5}	{1}	{2,3}
Wrkrs	-	{0}	{0}	{1}	-	-	-	-
Gndr	-	{1}	{2}	-	-	-	-	-
Fracc	-	-	-	-	{0}	{1}	-	-
Prov	-	-	-	-	-	-	-	{1,4,7,8,9,10,11,12}
Begin	.2174	.3073	.4375	.3396	.3608	.5970	.3191	.3581
Mid	.4203	.4203	.3611	.5472	.4433	.2910	.3830	.6157
End	.3623	.2724	.2014	.1132	.1959	.1120	.2979	.0262
N	69	179	144	53	97	134	47	229
Rule#	1	2	3	4	5	6	7	8

Strikingly, the most significant condition is the holiday region, where *people living the southern holiday region* of the Netherlands have higher probabilities of selecting the end, and in particular the mid part of the summer holiday (R14-24). This accords with the fact that, in 1998, schools in this part of the country were the last to close their doors for the summer holidays. Within the group of holidays made by people from the southern holiday region, the *number of holidays already scheduled for the second part of the summer holiday* is significant. As expected, when at least one holiday has already been scheduled for this part of the summer, the probability of scheduling additional holidays for the other two parts increases (R23+24). If no holidays have been scheduled for the second part of the summer holiday, the majority of holidays is scheduled for this period (R14-22), except when the holiday under consideration is shorter than 9 days (R14), or when this is an extended holiday made by a jobless tourist from a household with children aged 6-17 years who does not own a non-permanent tourist accommodation (R16).

Table 9.4 Continued - summer season period

Region	{1,2}	{1,2}	{1}	{2}	{1,2}	{3}	{3}	{3}
#Tspr	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	-	-	-
Chld	{2}	{2}	{2}	{2}	-	-	{1,3}	{2}
Lngth	{2,3}	{2,3}	{4,5}	{4,5}	-	{1,2}	{3}	{3}
Wrkrs	-	-	-	-	-	-	{0}	{0}
Fracc	-	-	-	-	-	-	-	{0}
Prov	{2,3,5}	{6}	-	-	-	-	-	-
#Ts2	-	-	-	-	-	(<-,0]	(<-,0]	(<-,0]
Begin	.2545	.6901	.6667	.8320	.2576	.1047	.2300	.0545
Mid	.5455	.2817	.3333	.1680	.3106	.4188	.4000	.4182
End	.2000	.0282	.0000	.0000	.4318	.4765	.3700	.5273
N	55	71	138	125	396	277	100	5273
Rule#	9	10	11	12	13	14	15	16

For holidays made by people living in the *northern and mid holiday region*s of the Netherlands, the number of holidays already scheduled for spring is important, where the mid and final period of the summer season gain importance when at least one holiday has already been scheduled for the spring season (R13). When *no holidays have been scheduled for the spring season*, the presence of children in the household is important. *Households with schoolchildren* (6-17 years) prefer the mid-summer season for short breaks, medium long holidays and extended holidays (R7-10; except for tourists originating from the Province of Gelderland (R10)), whereas the beginning of the summer holiday season is preferred for long and extra long holidays (R11+12). Overall, for these households, the final part of the summer holidays is not very popular. This applies especially to long and extra long holidays.

Table 9.4 Continued - summer season period

Region	{3}	{3}	{3}	{3}	{3}	{3}	{3}	{3}
#Tspr	-	-	-	-	-	-	(<-,1]	(1,->)
Chld	{2}	-	-	-	-	-	-	-
Lngth	{3}	{3}	{4,5}	{4,5}	{4,5}	{4,5}	-	-
Wrkrs	{0}	{1}	-	-	-	-	-	-
Fracc	{1}	-	{0}	{1}	{0}	{1}	-	-
#Ts2	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)
Party	-	-	{1,3,99}	{1,3,99}	{2,4,5}	{2,4,5}	-	-
Begin	.0270	.0526	.0377	.1059	.2000	.3939	.2586	.5738
Mid	.7297	.7632	.7170	.8118	.4444	.4394	.3276	.1475
End	.2433	.1842	.2453	.0823	.3556	.1667	.4138	.2787
N	37	38	53	85	45	66	58	61
Rule#	17	18	19	20	21	22	23	24

Finally, for *households without or with very young children only* (R1-6), the preferences for the three parts of the summer are rather indistinct, where the preference for the final part decreases with the duration of the holiday (but not as strong as for households with schoolchildren), and where the preferences for the first two parts for holidays of at least 5 days are determined by restrictions imposed by the employer, gender and the possession of a tourist accommodation without a permanent place.

9.3.4 The DT for the choice of departure date

The final DT for tourist timing choices represents the conditions under which people decide to depart on Friday or Saturday, when congenstion is more likely, or on other days of the week. Table 9.5 shows the 68 decision rules that were induced. In this table, “FriSat” indicates the probability of departing on Friday or Saturday under the specified conditions. In contrast, “Other”, indicates the probability of departing on Sunday, Monday, Tuesday, Wednesday or Thursday, the “off-peak” days of the week.

Table 9.5 DT for the choice of departure date

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
#Tfs	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]
Lngth	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]
Seas+	{1,3}	{6}	{1,3,6}	{1,3,6}	{2}	{2}	{2}	{2}	{2}	{4,5}
Acc	{1,3}	{1,3}	{2,4}	{2,4}	-	-	-	-	-	-
HolP	-	-	{0}	{1}	-	-	-	-	-	-
Chld	-	-	-	-	{1,2}	{1,2}	{1,2}	{1,2}	{3}	-
Tcost	-	-	-	-	(<-,518]	(518, 900)	(900, ->)	(900, ->)	-	-
Wrkhr	-	-	-	-	-	-	(<-,0]	(0,->)	-	-
Other	.2051	.4286	.0556	.1831	.4000	.1707	.5313	.2578	.1493	.5057
FriSat	.7949	.5714	.9444	.8169	.6000	.8293	.4687	.7422	.8507	.4943
N	78	77	126	71	125	82	64	128	67	176
Rule#	1	2	3	4	5	6	7	8	9	10

The most significant condition for the choice of departure date is the anticipated destination of the holiday in question, where domestic holidays are more likely to begin on a Friday or Saturday (65.1% vs. 47.3%). For *domestic holidays* (R1-37), the number of tourist trips that have already started on a Friday or Saturday is the next most significant condition. If people have already started *4 or more holidays on these two days*, they are very likely to depart on these days again for subsequent holidays (R 33-37).

Table 9.5 Continued - departure date

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
#Tfs	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]
Lngth	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]
Acc	-	-	{1,3}	{1,3}	{2,4}	{2,4}	{2,4}	{2,4}	{2,4}	{2,4}
HolP	{0}	{0}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Chld	{1,3}	{2}	-	-	-	-	-	-	-	-
Wrkhr	-	-	(<-,0]	(0,->)	-	-	-	-	-	-
Prio4	-	-	-	-	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,2]	(2,->)
HHsz	-	-	-	-	(<-,2]	(2,->)	(2,->)	(2,->)	-	-
Prov	-	-	-	-	-	{1,2,4, 8,10,12}	{3,5,6,7, 9,11}	{3,5,6,7, 9,11}	-	-
Gndr	-	-	-	-	-	-	{1}	{2}	-	-
Other	.6536	.9016	.5422	.8235	.5102	.4286	.2400	.0926	.4257	.5965
FriSat	.3464	.0984	.4578	.1765	.4898	.5714	.7600	.9074	.5743	.4035
N	332	61	83	51	49	70	50	54	202	57
Rule#	11	12	13	14	15	16	17	18	19	20

If *two or three holidays* have already started on Friday or Saturday, additional domestic tourist trips are most likely to begin on these days too (R28-32), except for holidays of at least 5 days during winter, spring or the last part of summer (R31).

Finally, if the tourist has started *at most one tourist trip on Friday or Saturday*, the duration of the domestic holiday under considerations is important in the choice of the departure date. Under these conditions, *short breaks* usually commence on Friday or Saturday (R1-10), except for short breaks starting during the last six weeks of the summer holidays (R10) and for short breaks during the spring pursued by a jobless tourist from a household with no or schoolgoing children, who has already spend more than NLG 900 on holidays (R7).

Table 9.5 Continued - departure date

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
#Tfs	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,3]	(1,3]	(1,3]
Lngth	(2,4]	(2,4]	(2,4]	(2,3]	(3,4]	(2,4]	(4,->)	(<-,1]	(<-,1]	(<-,1]
Seas+	-	-	{1,2}	{3,4,5,6}	{3,4,5,6}	-	-	-	-	-
Acc	{1,4}	{2,3}	-	-	-	-	-	-	-	-
HolP	-	-	-	-	-	-	-	{0}	{0}	{1}
HHsz	(<-,2]	(<-,2]	(2,4]	(2,4]	(2,4]	(4,->)	-	-	-	-
Work	-	-	-	-	-	-	-	{0,1,2}	{3}	-
Other	.4222	.6379	.4828	.1963	.3111	.1368	.4815	.1183	.3167	.3237
FriSat	.5778	.3621	.5172	.8037	.6889	.8632	.5185	.8817	.6833	.6763
N	90	116	58	163	90	117	54	169	60	241
Rule#	21	22	23	24	25	26	27	28	29	30

Table 9.5 Continued - departure date

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{2}	{2}	{2}
#Tfs	(1,3]	(1,3]	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	-	-	-
Lngth	(1,->)	(1,->)	-	-	-	-	-	(<-,1]	(<-,1]	(<-,1]
Seas+	{1,2,5}	{3,4,6}	-	-	-	-	-	{1,3,4,5,6}	{1,3,4,5,6}	{2}
HolP	-	-	-	-	-	-	-	{0}	{0}	{0}
Prov	-	-	{1,2,3,4,5, 6,8,10}	{7,9, 11,12}	-	-	-	-	-	-
Gndr	-	-	-	-	-	-	-	{1}	{2}	-
Sclss	-	-	(<-,1]	(<-,1]	(1,2]	(2,->)	(2,->)	-	-	-
#Tp	-	-	-	-	-	(<-,1]	(1,->)	-	-	-
WrkW	-	-	-	-	-	-	-	{0,1,2,3,4}	{0,1,2,3,4}	{0,1,2,3,4}
Other	.7701	.4464	.0000	.2083	.2759	.1633	.0327	.0727	.2200	.3188
FriSat	.2299	.5536	1.000	.7917	.7241	.8367	.9673	.9273	.7800	.6812
N	87	56	60	48	58	49	336	55	50	69
Rule#	31	32	33	34	35	36	37	38	39	40

At the other end of the continuum, *extra long holidays* are also more likely to start on a Friday or Saturday, but for this trip type differences between the two parts of the week are small (52% vs. 48%; R27). *Extended and long domestic holidays* too are more likely to start on Friday or Saturday (R21-26), except for holidays made by tourists from 1- and 2-person households who spend the nights in their own accommodation (permanent and non-permanent; R22). *Medium long domestic holidays*, finally, are holidays that are more likely to commence on the off-peak days of the week (R11-20). However, even for medium long domestic holidays there is a large group that is (slightly) more likely to leave on a Friday or Saturday (R16-19). In particular this applies to medium long holidays that are scheduled during a school holiday period, and that are the second most important trip within the tourist's trip program (R19), or that have the highest profiling priority within the trip program of people from households with at least 3 members (R16-18).

With regard to *foreign holidays*, the duration of the trip is very important to the decision whether or not to leave on Friday or Saturday. *Short breaks abroad* often start on Friday of Saturday (R38-45). Two exceptions, however, exist. First, when the tourist does not own a non-permanent accommodation, and (s)he has not yet started any holidays on Friday or Saturday, and the holiday under consideration is scheduled to take place during a school holiday period, a foreign short breaks is more likely to leave on an off-peak day of the week (R42). Secondly, this also applies to a short break during a school holiday period to France, the UK and Ireland or Scandinavia and Denmark by a tourist who owns a non-permanent tourist accommodation (R44).

Table 9.5 Continued - departure date

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
#Tfs	-	(<-,0]	(0,->)	-	-	-	-	-	-
Lngth	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,2]	(1,2]	(1,2]	(1,2]
HolP	{0}	{1}	{1}	{1}	{1}	{0}	{0}	{1}	{1}
Prov	-	-	-	-	-	-	-	{1,4,6,8,10,11}	{2,3,5,7,9,12}
WrkW	{9}	-	-	-	-	-	-	-	-
Fracc	-	{0}	{0}	{1}	{1}	-	-	-	-
D_A	-	-	-	{1,7,10}	{2,3,4,5, 6,8,9,11}	{1,2,6,7}	{3,11}	{1,2,3,6,7,11}	{1,2,3,6,7,11}
Other	.4359	.7018	.4455	.5686	.2564	.6567	.8889	.6821	.4355
FriSat	.5641	.2982	.5545	.4314	.7436	.3433	.1111	.3179	.5645
N	78	57	101	51	78	134	63	151	124
Rule#	41	42	43	44	45	46	47	48	49

For *medium long holidays abroad*, the destination is important. Medium long holidays to France, Spain and Portugal, Belgium and Luxembourg, Italy and Greece, the UK and Ireland and “other” destinations outside Europe, often leave on off-peak days of the week (R46-49), except for trips during school holidays by tourist from Friesland, Drente, Flevoland, Utrecht, South-Holland and Limburg (R49). Medium long holidays to Germany, former eastern European countries, south-east Mediterranean countries, Scandinavia and Denmark and “other Europe” are most likely to start on off-peak days of the week when they are scheduled outside school holiday periods (R50+51). During school holiday periods, however, medium long holidays to these countries are also slightly more likely to commence on Friday or Saturday (54% vs. 46%; R52). Medium long holidays to Austria and Switzerland, finally, are most likely to leave on Friday or Saturday (R53+54).

Table 9.5 Continued - departure date

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Lngth	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]
Seas+	-	-	-	-	-	{1}	{1}	{2}	{2}	{2}
HolP	{0}	{0}	{1}	-	-	{0}	{1}	-	-	-
Gndr	{1}	{2}	-	-	-	-	-	-	-	-
Fracc	-	-	-	{0}	{1}	-	-	-	-	-
D_A	{4,8,9,10}	{4,8,9,10}	{4,8,9,10}	{5}	{5}	-	-	-	-	-
Tmode	-	-	-	-	-	-	-	{1,3,4}	{1,3,4}	{2,5}
Age	-	-	-	-	-	-	-	(<-,48]	(48,->)	-
Other	.5238	.7451	.4606	.3443	.1739	.2706	.1096	.2500	.5930	.6718
FriSat	.4762	.2549	.5394	.6557	.8261	.7294	.8904	.7500	.4070	.3282
N	63	51	165	61	46	85	73	60	86	131
Rule#	50	51	52	53	54	55	56	57	58	59

For *extended holidays abroad*, the season is an important factor. Foreign holidays during the winter are very likely to start on Friday or Saturday (R55+56). During the spring season, however, the off-peak days of the week are more probable (R57-59), except for car-, train- or bus-based holiday by people younger than 49 years (R57). During the first part of the summer holiday too, extended holidays often start at off-peak days of the week (R60). During the remaining parts of the year, Friday and Saturday are the favourite days of departure for car-based holidays (R61), whereas the off-peak days are more likely for extended foreign holidays by other transport modes (R62+63). *Long and extra long holidays abroad*, finally, often depart on the off-peak days of the week (R64-68), except for holidays during the winter and the last part of the summer holiday season to France, Spain, Portugal, Belgium, Luxembourg, Greece, Italy, and Scandinavia or Denmark (R64).

Table 9.5 Continued - departure date

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Lngth	(2,3]	(2,3]	(2,3]	(2,3]	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)
Seas+	{3}	{4,5,6}	{4,5,6}	{4,5,6}	{1,5}	{2,3,4,6}	-	-	-
HHsz	-	-	-	-	-	-	(<-,1]	(1,3]	(3,->)
D_A	-	-	-	-	{1,2,3, 5,6,10}	{1,2,3, 5,6,10}	{4,7,8, 9,11}	{4,7,8, 9,11}	{4,7,8, 9,11}
Tmode	-	{1}	{2,3,4,5}	{2,3,4,5}	-	-	-	-	-
Cst	-	-	{1,5}	{2,3,4,6}	-	-	-	-	-
Other	.7025	.3941	.5113	.7586	.3649	.5957	.5800	.8198	.6494
FriSat	.2975	.6059	.4887	.2414	.6351	.4043	.4200	.1802	.3506
N	121	203	133	87	74	517	50	172	77
Rule#	60	61	62	63	64	65	66	67	68

9.4 Statistics and validation

To complete this chapter on tourist timing decisions, this final section discusses the most important statistics and the predictive abilities of the generated DTs. The confusion matrix and statistics for the DT for the choice of “during school holiday period or not” are presented in Table 9.6. Using a deterministic assignment rule, this DT would correctly classify 66.67% of all observations. This means only a small improvement compared to the null-model that would produce a 60.06% hit score. For the probabilistic DT, the improvement is slightly more important, from 52.03% for the null-model to 60.17% for the 46 decision rules. From the confusion matrix it can be learned that this increase should mainly be ascribed to the improved ability of the model to predict trips outside school holidays (20.36%), because the improvement for trips during school holidays is only 11.35% (from 60.06% to 66.85%). The moderate performance of the DT is confirmed by the value of the contingency coefficient, that is only 0.3812 on a maximum of 0.7071 for this 2×46 DT.

Table 9.6 Confusion matrix & statistics for the DT *school holiday period or not*

	Outside school holiday'	In school holiday'	Total'
Outside s. holiday	**.5015**	.4985	.3994
In school holiday	.3315	**.6685**	.6006
Total	.3994	.6006	**.6017**
Nº of observations	7121		
Stopping criteria	80 before/ 35 after		
α	0.05		
Nº of columns (= rules)	46		
Theta (1 col)	4277 (.6006)[a]	3705 (.5203)[b]	
Theta (46 cols)	4748 (.6667)[a]	4285 (.6017)[b]	
χ^2_{DT}(df) / Cont. coef.	1210.91(45) / 0.3812		

[a] Deterministic [b] Probabilistic

Table 9.7 presents the indicators for the DT for the choice of season. This DT performs much better than the first DT for timing decisions given an improvement of the predictive abilities of 80.82% (from 30.14% to 54.50% for the probabilistic DT; the deterministic DT improves by 78.57% to 65.50%).

Table 9.7 Confusion matrix & statistics for the DT *season*

	Winter'	Spring'	Summer'	Autumn'	Total'
Winter	**.2087**	.4112	.1791	.2010	.1188
Spring	.1363	**.5439**	.1381	.1817	.3582
Summer	.0581	.1334	**.7586**	.0500	.3668
Autumn	.1533	.4261	.1185	**.3021**	.1562
Total	. 1188	. 3582	.3668	.1562	**.5450**
Nº of observations	7121				
Stopping criteria	60 before/25 after				
α	0.05				
Nº of columns (= rules)	88				
Theta (1 col)	2612 (.3668)[a]	2146 (.3014)[b]			
Theta (88 cols)	4664 (.6550)[a]	3881 (.5450)[b]			
χ^2_{DT}(df) / Cont. coef.	5792.07(261) / 0.6696				

[a] Deterministic [b] Probabilistic

This is supported by the value of the contingency coefficient that is rather high (0.6696 out of a maximum 0.8660 for this 4×88 DT). Evidently, an important part of this improvement should be attributed to the condition variable representing the first timing decision, because all trips scheduled during the summer are taken during the school holiday period. The confusion matrix shows that the percentages of correctly classified observations nearly double for all seasons with the exception of spring (from 35.84% to 54.39% - still a very important improvement).

Table 9.8 Confusion matrix & statistics of the DT *summer season period*

	Begin'	Mid'	End'	Total'
Begin	**.4619**	.3615	.1766	.3361
Mid	.2924	**.4702**	.2374	.4142
End	.2393	.3974	**.3633**	.2496
Total	.3361	.4142	.2496	**.4407**

N° of observations	2612	
Stopping criteria	80 before/35 after	
α	0.05	
N° of columns (= rules)	24	
Theta (1 col)	1082 (.4142)[a]	906 (.3467)[b]
Theta (24 cols)	1393 (.5333)[a]	1151 (.4407)[b]
χ^2_{DT}(df) / Cont. coef.	2300.05 (46) / 0.4765	

[a] Deterministic [b] Probabilistic

Table 9.8 presents the statistics and validity indicators of the DT that describes the choice between the three parts of the summer season. Compared to the null-model, the 24 decision rules predict these choices 27.04% more accurately (28.74% in case of deterministic assignment). This is mainly due to the increased number of hits for the first and the last part of the summer because for the middle part, the model only increases the predictive accuracy from 41.42% to 47.02%. Also, the contingency coefficient is rather poor at 0.4762 out of a maximum of 0.8165 for this 3×24 DT.

Table 9.9 Confusion matrix & statistics for the DT *departure date*

	Fri- or Saturday'	Other day'	Total'
Fri- or Saturday	**.6584**	.3416	.5693
Other day	.4515	**.5485**	.4307
Total	.3994	.6006	**.6110**

N° of observations	7121	
Stopping criteria	100 before/ 45 after	
α	0.05	
N° of columns (= rules)	68	
Theta (1 col)	4054 (.5693)[a]	3629 (.5096)[b]
Theta (68 cols)	4944 (.6943)[a]	4351 (.6110)[b]
χ^2_{DT}(df) / Cont. coef.	1473.65(67) / 0.4141	

[a] Deterministic [b] Probabilistic

Finally, the statistics and validation results of the last timing decision are presented in Table 9.9. Again the improvement in predictive accuracy and the contingency coefficient appear to be rather poor. Nevertheless, the 68 decision rules are still able to correctly predict 61.10% of the observations against 50.96% by the null-model, an improvement of 19.90%. The improvement is most notable for the category “other days of the week” (from 43.07% to 54.85%). The contingency coefficient of this 2×68 DT is 0.4141, where the maximum is 0.7071.

9.5 Conclusion and discussion

This chapter described three dimensions of tourist timing decisions that are important to the tourism industry. These dimensions include the decision to (1) travel during a school holiday period or not; (2) select a particular season; and (3) leave on a particular day of the week in terms of “peak” (Friday and Saturday) and “off-peak” days. The choice of season was described using two DTs, the first of which captures the choice of season in terms of the four seasons winter, spring, summer and autumn. Conditional upon the choice for a summer holiday, a second DT described the choice for a particular part of this season (begin, mid or end). The first two dimensions are considered following the choice of travel party, and their mutual profiling priority was determined on empirical grounds. Based on conceptual considerations, however, the third dimension is considered only when the two other timing dimensions and the choice of destination, accommodation and transport mode have been decided upon. The next chapter describes the profiling decision following these timing decisions, i.e. the choice of destination.

10 Destination Choices

This chapter discusses the induction and representation of tourist destination choices. Several stages in the destination choice process are distinguished, the first of which concerns the choice between domestic and foreign destinations. In total, eight decision tables represent a hierarchy of destination choice sets. Also, a link between destination and transport mode choices is established by including the individual tourist's "general propensity to select particular transport modes" as conditions in the choice of destination.

10.1 Introduction

Given the choice of travel party and the timing of the tourist trip, the next step in the *MERLIN*-system is to determine the destination. Based on the conceptual considerations discussed in chapter 2, there are two important issues related to destination choice, including the conceptualisation of these choice processes and their relationship to transport mode choices. With regard to the first issue, the phased nature of the tourist decision-making process and the choice sets structure are core to the conceptualisation of the tourist's destination choices (Um & Crompton, 1990; Woodside & Lysonski, 1989). Potential tourist destinations are systematically excluded based on the traveller's personal and household constraints and preferences and on the destination characteristics. The alternatives in the destination choice sets of the *MERLIN*-system will be based on regional classifications at different scale levels.

With regard to the second issue, it has been suggested that the choices of destination and transport mode are interrelated. Dellaert (1995) found that combined destination-transport choices can be targeted most effectively through the destination choice. For this reason it was decided to model the mode choice (see chapter 11) conditional upon the destination choice. Intuitively, however, one would assume changes in the transport system and the subsequent mode choices to affect destination choices. A significant change in the costs of flying, for instance, might induce people to refrain from flying and/or to select destinations closer to home. The *MERLIN*-system will therefore pay special attention to the interrelationship between destination and transport mode choices.

Given these two issues, the purpose of this chapter twofold. First, the subsequent section discusses how transport mode preferences are represented in the destination choice models. Next, a hierarchy of tourist destination choice sets is presented. For each of these choice sets a decision table will be induced using the same approach as in the previous chapters. Finally, summarising the major findings with regard to tourist destination choices will conclude this chapter.

10.2 General propensities to select transport mode alternatives

The *MERCIN*-system assumes destination and transport mode choices to be interrelated. And, since destination choices are considered prior to transport modes, destination choice decision rules in the system should be able to 'look ahead' in order that changes in the transport system may affect destination choices. This objective is attained by including condition variables that express the tourist's general propensities to select the most important transport mode alternatives for holidays, i.e. "car", "aeroplane" or "alternative" (on a scale from 0 to 100). Each variable expresses the *predicted individual propensity* to select one of these three mode alternatives and is calculated using a MNL-model. First, the observed mode preferences for each individual are calculated based on the CVO-data. E.g., a tourist with 2 car-based holidays and 1 holiday by air would have:

- Observed individual "car" preference: (2/3 * 100) = 66.7
- Observed individual "aeroplane" preference: (1/3 * 100) = 33.3
- Observed individual "alternative" preference: (0/3 * 100) = 0.00

The MNL-model is calibrated relating the observed individual mode preferences to (1) the aggregate marginal distribution of the transport mode alternatives over all travellers and over all observed holidays[30]; (2) the individual's annual number of holidays categorised into 3 levels: one holiday per year (#holidays: 1), 2-3 holidays per year (#holidays: 2-3) and 4 or more holidays (reference); (3) the socio-economic conditions presented in table 7.1; and (4) the following additional personal and household conditions (all levels are dummy coded relative to the last/highest level):

- age 0-15 years (relative to the age-category: 56 years and older);
- the possession of skis; and
- the holiday region in the Netherlands: North (Holiday region: North), Mid (Holiday region: Mid) or South (reference).

The "reference-person" thus constitutes an elderly (56+ years) male without children in the household, without a car or (permanent or free) tourist accommodation, with higher levels of income and education and living in a rural area in the south of the country, who does not have a paid job and pursues at least 4 holidays a year. Note that these conditions are related to the *individual traveller* rather than the observed holidays.

With the exception of the aggregate marginal distribution of the mode alternatives, all conditions are entered as alternative specific parameters. Using

[30] The distribution over all travellers and over all observed holidays in the 1998 CVO-data set amounted to 74.9 ("car"), 12.2 ("aeroplane") and 12.9 ("other").

dummy coding, "other" is selected as the reference alternative. The parameters that did not prove to be significant at the 95% level of confidence are excluded. The model parameters are presented in Table 10.1. This model thus includes 38 parameters (LL(0) = -260,979.45 and LL(B) = -153,999.52) and performs reasonably well as indicated by the Rho-squared (AIC) = .410.

Table 10.1 Parameter estimates for the individual propensity to select a transport mode relative to the "other mode"-alternative; *t*-values in brackets

	CAR	AIR
Marginal distribution modes	.0118 (29.6)	
annual number of holidays		
#holidays: 1	-.336 (-19.0)	.379 (17.5)
#holidays: 2-3	-.278 (-16.6)	.048 (2.30)
socio-demographic characteristics		
Income < NLG 35,000	-.117 (-7.15)	-.520 (-26.4)
Income NLG 35-55,000	.150 (10.5)	-.271 (-15.4)
Education low	-.155 (-8.84)	-.300 (-14.1)
Education intermediate	-.082 (-4.86)	-.179 (-9.03)
Age 0-15 years	.341 (14.2)	.293 (8.68)
Age 16- 30 years	-.597 (-29.6)	.078 (3.08)
Age 31-55 years	-.073 (-3.73)	-.311 (12.8)
Female		.168 (13.9)
Household size	.059 (9.00)	-.245 (-28.8)
Child in household	.392 (20.5)	-.222 (-8.97)
Urban residence	-.236 (-25.5)	
At least one car	.990 (69.4)	.611 (34.5)
Permanent accommodation	-.040 (-1.92)	-1.09 (-31.1)
Accommodation without place	.322 (25.9)	-.286 (-17.6)
Work: employed	.053 (2.78)	.061 (2.52)
Working hours per week	.005 (10.0)	.013 (18.8)
Skis		
Holiday region: North	-.114 (-9.78)	.226 (14.3)
Holiday region: Mid		.220 (15.7)

The parameter for the aggregate marginal distribution of the transport mode alternatives (.0118) has the expected sign, indicating the propensities of a "reference-person" to travel by car, air or an alternative mode are positively related to the aggregate marginal distribution of transport modes. With regard to the socio-economic condition variables it can be concluded that relative to the alternative modes and the "reference-person", the *propensity to travel by car* is positively affected by the intermediate income level (.150), being a child (Age 0-15 years: .341), household size (.059), the presence of children (.392) and/or (a) car(s) in the household (.990), having a paid job (.053) and the number of working hours per week (.005) and the possession of a tourist accommodation without a permanent place (.322). On the other hand, the *propensity to travel by air* relative to the “alternative” option and the “reference-person”, is positively affected by lower number of holidays per year (.379 and .048), being a child (Age 0-15 years: .293), in your late teens or twenties (Age 16-30: .078), a woman (.168), owning (a) car(s) (.611), having a paid job (.061) and more working hours per week (.013), and living in the Northern (.226) or Mid (.220) holiday region.

Given these model parameters, the general propensities to select each of the three available transport mode alternatives (on a scale from 0 to 100) are predicted for each person in the sample, where the sum of these propensities always sums up to 100. Finally, the three “propensity”-scores for each individual are assigned to the holidays made by this person. These variables are entered as conditions into the eight decision tables for tourist destination choices.

10.3 The question of where to travel

10.3.1 The choice sets

The *MERLIN*-system considers destination choice sets that are based on regional classifications at different scale levels. Three types of tourist destination regions can be distinguished (Dietvorst, 1993c, p.67-68). First, *a-priori regions* are defined by governments or by the tourism industry. Examples at different scale levels include countries, municipalities, provinces and Designated Outdoor Recreation Areas. Secondly, *homogeneous regions* are defined based on a set of objective, internal similarities such as regional characteristics. An example of a homogeneous regionalisation in the Netherlands are the 17 “tourist areas” identified by the Statistics Netherlands (Dutch: *CBS*). Third and finally, *functional regions* are based on high degrees of internal interaction. Tourist-recreation complexes are examples of functional regions. Obviously, a-priori, homogeneous and functional regions can overlap each other. *MERLIN* identifies regions based on their relevance to planning agencies (mainly a-priori regions) as well as to the experience and perception of tourists (homogeneous or functional regions).

In a relatively small country like the Netherlands, one of the most important stages of the destination choice comprises the decision whether or not to travel abroad. Based on a longitudinal analysis of holiday patterns, for instance, Jansen-Verbeke and Spee (1995) concluded that the Dutch market can be divided into two segments. The first group of people tends to spend the most important holiday in the Netherlands, using their own car, organising their holiday themselves, and spending, on average, less money than the second group of people, who focus on foreign destination (Jansen-Verbeke & Spee, 1995, p. 50). Similarly, the NRIT used a classification system that determines whether a Dutch tourist is likely to select domestic destinations for both main and extra holidays (type: "plus-plus"), foreign destinations both main and extra holidays (type: "min-min"), domestic destinations for main and foreign destinations for extra holidays (type: "plus-min"), and foreign destinations for main and domestic destinations for extra holidays (type: "min-plus") (Dirven *et al.*, 1998). The first destination choice set in the *MERLIN*-system therefore comprises the a-priori regions "domestic" and "abroad".

With regard to foreign holidays, *distance* appears to be an important factor in the decision making process. Schmidhauser (1976) identified four destination zones for Swiss residents, including "Switzerland" (domestic), "neighbouring countries", "other Europe" and "outside Europe" (Schmidhauser, 1976). Although not explicitly, this classification appears to be a mixture of geographical and cultural distance factors (Oppermann, 1998, p. 323).

Clearly, some foreign counties have competitive advantages in attracting Dutch tourists because of the geographical proximity that reduces travel costs and efforts. Similarly, cultural proximity reduces the risk of unpleasant experiences (language, food, social manners and so on). Conversely, cultural and psychological distance may attract more experienced and/or risk-pursuing tourists, and it may offer the prospect of a more attractive climate. Within the foreign destinations, the *MERLIN*-system therefore distinguishes between (geographically and/or culturally) neighbouring countries including France, Belgium/Luxembourg, Germany and the UK/Ireland, and more distant countries. The latter choice set is further detailed by distinguishing between the following (groups of) countries: Spain and Portugal; Austria and Switzerland; Italy and Greece; (former) Eastern European countries including the (former) Soviet Union, Yugoslavia, Czechoslovakia, Hungary, Bulgaria, Rumania, and Poland; South East Mediterranean countries such as Malt, Turkey, Morocco, Cyprus and Tunisia; Scandinavia and Denmark; and "Other more distant countries". The latter category includes all destinations outside Europe that are not part of any of the other categories.

With regard to domestic tourist trips, distance is probably less relevant to the destination choice process. Instead, the prospect of particular experiences offered by the regionally and locally available tourist-recreation facilities and services are more likely to play a part in domestic destination choices[31] (homogenous or functional regions). The previously mentioned Dutch tourist areas are mainly based on geographical and landscape characteristics, thus bearing a strong relationship to the possible tourist experiences. The Netherlands include the following tourist areas according to this classification, (see Figure 10.1):

1 The Wadden Islands;
2 The North Sea Coastal Area;
3 The beaches of the Lake Yssel;
4 The Delta area;
5 The lake areas in Groningen, Friesland and North-West Overijssel;
6 The lake areas of Holland and Utrecht;
7 The Utrecht elevations and Gooi-area;
8 The Veluwe and its borders;
9 The river area of Gelderland;
10 The Achterhoek;
11 Twente, Salland and Vecht-area;
12 The sand areas of Groningen, Friesland and Drenthe;
13 The Western and Mid parts of Brabant;
14 The Eastern part of Brabant, Northern and Mid parts of Limburg and realm of Nijmegen;
15 The Southern part of Limburg;
16 Amsterdam, Rotterdam, The Hague and Utrecht (cities); and
17 The remaining areas.

The disadvantages of this classification are twofold. First, the regions are not directly related to tourist planning agencies. In contrast, the 12 Dutch provinces, for instance, are also well-known entities to Dutch tourists. In addition, these provinces are equipped to develop and co-ordinate planning for recreation and tourism, both within the provincial authorities and in the provincial tourist information centres (TICs). In recent years, however, in addition to the provincial and local TICs, regional TICs have been established around (parts of) the CBS-tourist areas (e.g., South-Limburg, South-West Veluwe and Mid and South-West Friesland).

The second disadvantage is the high number of regions. This problem is solved by grouping kindred regions together into higher-order regions.

[31] Evidently the prospect of particular experiences also plays an important part in the selection of *foreign* destinations. However, in the case of foreign destinations distance appears to be more important because in addition to, or as a consequence of raising cultural, financial or psychological barriers, it also more or less approximates to the prospect of particular experiences.

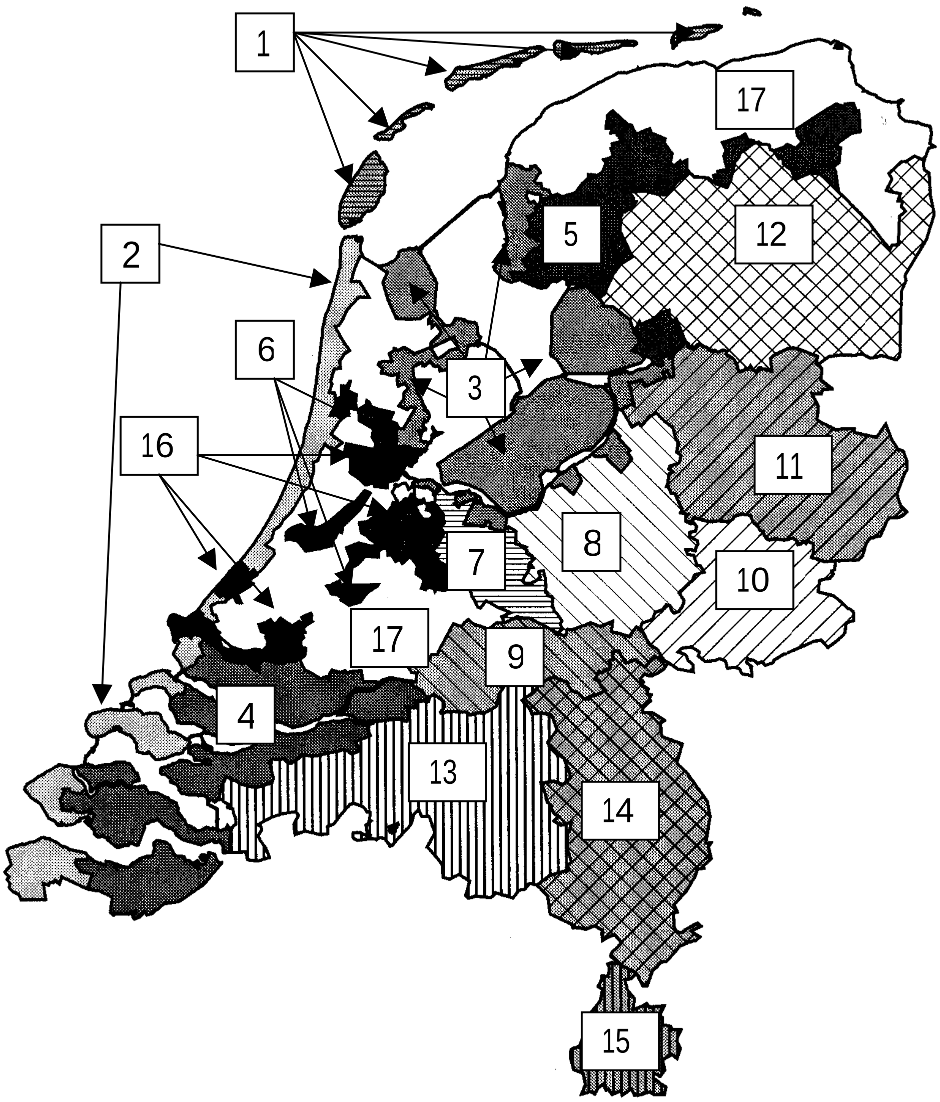


Figure 10.1 Dutch tourist areas (source: Statistics Netherlands)

Based on the above considerations, the 17 tourist areas are first reorganised into 9 regions. Next, these 9 regions are grouped together into a choice set of four “metaregions”; in this process, both the (dis)similarity between regions and the number of visits to each region were taken into account (the numbers between brackets refer to the original numbers in the CBS-classification; see Figure 10.1):

Metaregion WATER

(1,3,5) North: Water areas (beaches, lakes) in the North

(4,6,9) Mid: Water areas in the middle of the Netherlands

(2) N-Sea: The North Sea Coastal Area

Metaregion LAND NORTH
(7,8) Mid: Utrecht, ‘t Gooi and the Veluwe
(10,11) East: Achterhoek, Twente, Salland and the Vecht-area
(12) North: Sand areas in the North

Metaregion LAND SOUTH
(13) West: West & Mid Brabant
(14,15) South: East Brabant, Nijmegen & Limburg

Metaregion OTHER
(16,17) City: Cities in the Randstad Area and “other”

The resulting hierarchy of destination choice sets is shown in Figure 10.2. Each frame in this figure represents a DT. The following sections discuss the eight destination choice DTs in the MERLIN-system.

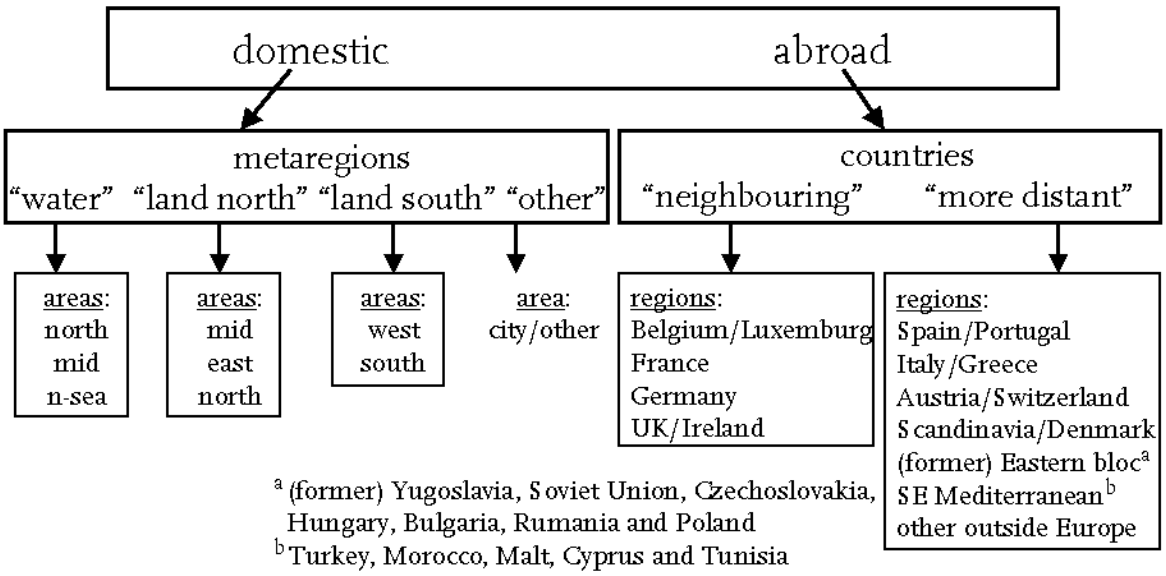


Figure 10.2 Phases in the tourist’s destination choice

10.3.2 The condition variables

As always, the condition variables common to all choice facets (Table A1.1 in Appendix 1) are entered into the CHAID-based algorithms for tourist destination choices. Also, the choice facets that have been considered previously, i.e. the travel party, the decision whether or not to travel during a school holiday period and the season, and the three conditions related to the profiling priority of the holiday are assumed to condition tourist destination choices (see Table A1.2). In addition, there are four condition variables specific to all destination choices (Table A1.7). First, based on the MNL-model that was discussed on section 10.2, there are three variables that represent the tourist's propensity to select the respective transport mode alternatives, including car, aeroplane and the alternative modes. The three propensities are each expressed on a scale from 0 to 100 and, save discrepancies due to rounding off, together always add up to 100.

The fourth condition variable common to the destination (and all subsequent) choices represent financial constraints in the tourist decision making process. *MERLIN* assumes that if, for instance, a lot of financial resources have already been allocated to tourist trips, people might be inclined to select cheaper destinations (or the other way around when people are generally inclined to select expensive holiday destinations). Hence, the total amount of expenditures that have already been spent on tourist trips with higher profiling priorities may condition subsequent choices. In the DTs this decision strategy can be captured by the variable Tcost, i.e. the summary variable of the "expenditures" that is updated as the profiling process for the tourist trips of an individual advances.

In addition to the variables that are common to all destination choices, each DT has a number of condition variables that are specific to the decision at that stage of the destination choice process (see tables A1.8 through A1.15). Basically, these conditions represent destination choices on trips with higher profiling priorities within the tourist trip program, both in terms of the number of trips to particular regions and in terms of the number of days associated with these trips. In the DTs for the final destination choices[32] the latter condition variables (i.e., the number of days already planned for the various regions) are not included to reduce the number of condition variables; this is particularly important because these DTs are not induced using all 7121 observed holidays. Finally, in most DTs, summary variables representing higher level destination choices on trips with higher profiling priorities are included (e.g., in the DT for the choice of the four Dutch metaregions, two condition variables represent the number of trips already scheduled in neighbouring and more distant foreign countries). The inclusion of these condition variables was based both on conceptual consideration and on empirical results (i.e., running the CHAID-based algorithm with different sets of conditions). In the end, however, the CHAID-based algorithm determines whether or not a condition variable will be included in the decision rules.

[32] I.e., the selection of tourist areas within the Dutch metaregions, the selection of a country given the decision to visit one of the neighbouring countries, and the selection of a particular destination zone given the decision to visit a more distant country.

10.3.3 The stopping criteria of the CHAID-based algorithm

The final decision regarding the DTs for tourist destination choices comprises the selection of the stopping criteria of the CHAID-based algorithm. For each destination DT, the α-level for predictor eligibility is set at 5%. With regard to the minimum number of observations before or after splitting decision rule, the performance of various models is reported in Appendix 2 (Tables A2.6 through A2.13). Based on these analysis, the stopping criteria for the choice of domestic or foreign destinations and the choice between neighbouring and more distant countries are set at 80 before and 35 after division because the resulting models perform best on the cross-validation measure, while the models with less decision rules (100 before/45 after split) perform worse at the aggregate level. For similar reasons, the DT for the choice of more distant destinations and the DT for the choice of neighbouring countries are induced with a CHAID-based algorithm that only allows splitting of decision rules comprising at least 35 observations before and 15 after the division.

With regard to the DTs for domestic destinations, the CHAID-based algorithm for the choice between the four Dutch metaregions performs best on the cross validation measure using very strict stopping criteria (100 before/45 after split), while the aggregate performance does not decreases compared to the other models. For similar reasons, the DTs for the tourist areas within the Dutch metaregions "Water" and "Land North" will be set at 60 observations before and 25 after splitting[33]. The stopping criteria for the choice of tourist areas within the Dutch metaregion "Land South", finally, is set at 60 observations before and 25 after splitting because this model performs best on the cross-validation measure.

10.4 The generated decision tables

This section discusses the structures of the eight DTs for destination choices generated by the CHAID-based algorithms. The first DT represents the choice between domestic and foreign destinations. Next, three DTs describe the choices between destinations abroad. The last four DTs, finally, represent destination choices within the Netherlands. Again, the discussion is restricted to the most important and/or most striking variables that condition tourist destination choices. Likewise, rules (Rule# in the DT), are indicated by R and the rule number.

[33] Stricter stopping criteria were not examined because of the low number of observations; this also applies to the selection of stopping criteria for the DTs for more distant destinations and for neighbouring countries (Tables A2.8 and A2.9).

10.4.1 The DT for the choice of domestic or abroad

The DT for the choice between domestic and foreign destinations was induced using all 7121 observed holidays. The resulting 74-column DT is presented in Table 10.2. In this table, “NL” denotes a domestic holiday and “Abroad” denotes a foreign holiday.

Table 10.2 DT for the choice of domestic (NL) or abroad

Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
#DD	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Lngth	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,2]	(1,2]	(1,2]	(1,2]
Age	(<-,36]	(<-,36]	(36,53]	(36,53]	(53,->)	(53,->)	-	-	-	-
Seas+	{1,3}	{2,4,5,6}	-	-	{1,4,5,6}	{2,3}	{1}	{2,5,6}	{2,5,6}	{2,5,6}
Gndr	-	-	{1}	{2}	-	-	-	-	-	-
Work	-	-	-	-	-	-	{0,1,2}	{0,1,2}	{0,1,2}	{0,1,2}
Tcost	-	-	-	-	-	-	-	(<-,2031]	(<-,2031]	(2031,->)
Party	-	-	-	-	-	-	-	{1,2,99}	{3,4,5}	-
NL	.3864	.6939	.6479	.4186	.7069	.9762	.1852	.3491	.7143	.1728
Abroad	.6136	.3061	.3521	.5814	.2931	.0238	.8148	.6509	.2857	.8272
N	44	147	71	86	58	42	81	169	35	81
Rule#	1	2	3	4	5	6	7	8	9	10

The most important condition variable in the choice between domestic and foreign destinations is the *possession of an accommodation with a permanent location*. For people owning such a tourist commodity, domestic destinations are often preferred (78.95% of the holidays made by these people are domestic; R61-74).

Table 10.2 Continued - domestic (NL) or abroad

Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
#DD	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Lngth	(1,2]	(1,2]	(1,2]	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]
Age	-	-	-	-	-	-	-	-	(<-,42]	(42,53]	(53,->)
Seas+	{3,4}	-	-	{1,5,6}	{1,5,6}	{1,5,6}	{2}	{2}	{3,4}	{3,4}	{3,4}
Work	{0,1,2}	{3}	{3}	{0,1}	{2,3}	-	{0,1}	{2,3}	-	-	-
Ski	-	{0}	{1}	{0}	{0}	{1}	-	-	-	-	-
NL	.5417	.5964	.2917	.0342	.1461	.0000	.0532	.2419	.1250	.3200	.5400
Abroad	.4583	.4036	.7083	.9658	.8539	1.000	.9468	.7581	.8750	.6800	.4600
N	72	223	48	117	89	62	94	124	88	50	50
Rule#	11	12	13	14	15	16	17	18	19	20	21

Table 10.2 Continued - domestic (NL) or abroad

Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
#DD	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,21]
Lngth	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(<-,1]
Age	(<-,61]	(<-,61]	(<-,61]	(<-,61]	(<-,61]	(<-,61]	(61,->)	(61,->)	-
Seas+	-	-	-	-	-	-	{1,6}	{2,3,4,5}	-
Prov	{1,2,3, 5,10}	{4,6,7,8, 9,11,12}	{4,6,7,8, 9,11,12}	{4,6,7,8, 9,11,12}	{4,6,7,8, 9,11,12}	{4,6,7,8, 9,11,12}	-	-	-
Palt[a]	-	(<- ,9.232]	(9.232, 10.369]	(10.369, 13.141]	(13.141, 17.566]	(17.566, - >)	-	-	-
NL	.1961	.0000	.1429	.0139	.0674	.0000	.0227	.2963	.7416
Abroad	.8039	1.000	.8571	.9861	.9326	1.000	.9773	.7037	.2584
N	51	35	35	72	83	81	44	81	418
Rule#	22	23	24	25	26	27	28	29	30

[a] Rounded off at 3 decimal fractions

This is probably because these accommodations are often in the Netherlands. There are, however, some exceptions. First, when the travel party is “alone” or “adults only” (and people do not own skis), foreign destinations become more attractive (R61+62). Also, people in the two highest social classes (without skis) who travel in large parties, or parties with children, are also more likely to select foreign destinations (R63). Furthermore, if people have already made at least two holidays abroad, selecting another foreign destination becomes more likely (R67). Presumably, this decision rule applies to people whose tourist accommodation is located abroad. Finally, people owning skis with no or only one previous domestic holiday also prefer foreign destinations more often (R72+73).

Table 10.2 Continued - domestic (NL) or abroad

Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Chld	{1}	{1}	{1}	{2,3}	{2,3}	{2,3}	{2,3}	{2}	{3}	{2,3}
#DD	(0,21]	(0,21]	(21,->)	-	-	-	-	-	-	-
Lngth	(1,->)	(1,->)	-	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,2]	(1,2]	(1,2]
Seas+	-	-	-	-	-	-	-	{1,2,6}	{1,2,6}	{3,5}
Party	-	-	-	{1,4,5,99}	{2,3}	{2,3}	-	-	-	-
Edu	(<-,4]	(4,->)	-	-	-	-	-	-	-	-
#TD	-	-	-	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Region	-	-	-	{1,2}	{1}	{2}	{3}	-	-	-
NL	.7159	.4921	.8229	.9000	.7143	.5122	.5920	.4549	.6639	.6460
Abroad	.2841	.5079	.1771	.1000	.2857	.4878	.4080	.5451	.3361	.3540
N	88	126	96	90	63	41	125	244	119	113
Rule#	31	32	33	34	35	36	37	38	39	40

Table 10.2 Continued - domestic (NL) or abroad

Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Chld	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}
Lngth	(1,2]	(1,2]	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]	(3,->)	(3,->)	(3,->)
Seas+	{4}	{4}	{1,2,6}	{3,4,5}	{3,4,5}	{3,4,5}	{3,4,5}	-	-	-
Gndr	{1}	{2}	-	-	-	-	-	-	-	-
Edu	-	-	-	-	(<-,3]	(3,4]	(4,->)	-	-	-
#TD	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
WrkW	-	-	-	(<-,1]	(1,->) + {9}	(1,->) + {9}	(1,->) + {9}	-	-	-
Inc	-	-	-	-	-	-	-	(<-,3]	(3,5]	(3,5]
#Tp	-	-	-	-	-	-	-	-	(<-,2]	(2,->)
NL	.8929	.7377	.1942	.5959	.3077	.7143	.3030	.3860	.1733	.5000
Abroad	.1071	.2623	.8058	.4041	.6923	.2857	.6970	.6140	.8267	.5000
N	56	61	139	245	39	35	66	57	75	36
Rule#	41	42	43	44	45	46	47	48	49	50

For people who *do not own a tourist accommodation with a permanent location*, the next most significant condition variable is the presence of children in the household. For *childless households* (R1-33), the number of tourist days already spent the Netherlands is important. If this number exceeds 21, the Netherlands is the preferred destination (R33). This decision rule indicates a strong preference for domestic holidays. In contrast, if the number of days already spent in the Netherlands does not exceed 21, but at the same time is larger than 0, foreign destinations are preferred (R30+31) except for holidays of a minimum 5 days made by higher educated people (R32).

Table 10.2 Continued - domestic (NL) or abroad

Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Chld	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}
Lngth	(3,->)	(3,->)	(3,->)	-	-	-	-	-
Age	(<-,8]	(8,->)	-	-	-	-	-	-
Tcost	-	-	-	(<-,900]	(<-,900]	(900,->)	(900,->)	(900,->)
Prov	-	-	-	{1,4,5,6,7,8,10}	{2,3,9,11,12}	-	-	-
#TD	(<-,0]	(<-,0]	(<-,0]	(0,1]	(0,1]	(0,1]	(0,1]	(0,1]
Inc	(5,7]	(5,7]	(7,->)	-	-	-	-	-
Freed	-	-	-	-	-	(<-,98]	(98,->)	(98,->)
HolP	-	-	-	-	-	-	{0}	{1}
NL	.3889	.1269	.0763	.8790	.6832	.3810	.8293	.6200
Abroad	.6111	.8731	.9237	.1210	.3168	.6190	.1707	.3800
N	36	134	118	157	161	63	41	50
Rule#	51	52	53	54	55	56	57	58

For those childless households without a permanent tourist accommodation who have not yet spent any tourist days in the Netherlands, foreign destinations are preferred, in particular for extended, long and extra long holidays (R14-29; except for extended holidays made by older tourists (> 53 years) during the first six week of the summer season (R21)). Shorter breaks, on the other hand, are more likely to be spent in the Netherlands (R1-6), although there are some notable exceptions (R1+4). Medium long holidays, finally, sit in-between (R7-13). In the latter case, the working situation, the possession of skis, the season, the travel party and the total amount of expenditures up to then are important.

Table 10.2 Continued - domestic (NL) or abroad

Peracc	{0}	{0}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Chld	{2,3}	{2,3}	-	-	-	-	-	-	-
Party	-	-	{1,2}	{1,2}	{3,4,5}	{3,4,5}	{99}	{99}	{99}
Ski	-	-	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Prov	-	-	-	-	-	-	{1,4,11,12}	{2,3,5,6,7,8,9,10}	-
#TD	(1,->)	(1,->)	(<-,0]	(0,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]
Inc	(<-,4]	(4,->)	-	-	-	-	-	-	-
Sclss	-	-	-	-	(<-,2]	(2,->)	-	-	-
#AA	-	-	-	-	-	-	(<-,1]	(<-,1]	(1,->)
NL	1.000	.8471	.3250	.5741	.4026	.6912	.7632	.9823	.3714
Abroad	.0000	.1529	.6750	.4259	.5974	.3088	.2368	.0177	.6286
N	60	170	120	54	77	68	38	113	53
Rule#	59	60	61	62	63	64	65	66	67

For *households with children who do not own a tourist accommodation with a permanent location* (R34-60), the number of prior domestic holidays is the most significant condition. For those who have already spent one or more holidays in the Netherlands, domestic holidays are preferred (R54-60; except when people have already spent at least NLG 900 on tourist trips and when they do not have missing values on the number of paid holidays in 1998 (R56)). These decision rules indicate very strong preferences for domestic holidays.

Alternatively, for tourists from similar households who have not spent any holidays in the Netherlands yet (R34-53), the duration is very important. Under these conditions, short breaks and medium long holidays are preferably spent in the Netherlands (R34-42; except for medium long holidays by households with schoolgoing children (R38)). For long and extra long holidays, on the other hand, foreign destinations appear to be more attractive (R48-53). These decision rules indicate people with very strong preferences for foreign destinations for longer holidays to spend their shorter holidays in the Netherlands.

Table 10.2 Continued - domestic (NL) or abroad

Peracc	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Party	{1,2,3,5}	{4,99}	{4,99}	-	-	-	-
Ski	{0}	{0}	{0}	{0}	{1}	{1}	{1}
#TD	(1,4]	(1,4]	(1,4]	(4,->)	(<-,1]	(<-,1]	(1,->)
WrkW	-	-	-	-	(<-,2]	(2,->) + {9}	-
Wrkhr	-	(<-,36]	(36,->)	-	-	-	-
NL	.7627	1.000	.9444	.9962	.4138	.1562	1.000
Abroad	.2373	.0000	.0556	.0038	.5862	.8438	.0000
N	59	186	36	522	58	64	43
Rule#	68	69	70	71	72	73	74

10.4.2 The DT for the choice of neighbouring or more distant countries

Given the choice to go abroad, the next stage in the destination choice process, is to decide whether to visit neighbouring countries, including France, Germany, Belgium/Luxembourg and the UK/Ireland, or more distant destinations. Using the 3266 observed foreign holidays, the CHAID-based algorithm induced a set of 34 exclusive and exhaustive decision rules that are presented in Table 10.3. In this table, “Distant” indicates the probability of selecting a more distant tourist destination, whereas the row labelled “Neighb” indicates the probability of selecting a destination closer to the Netherlands.

Table 10.3 DT for the choice between neighbouring and more distant countries

#Drem	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Lngth	(<-,1]	(1,2]	(1,2]	(1,2]	(2,4]	(2,4]	(2,4]	(2,4]	(2,4]
Wrkhr	-	(<-,0]	(0,20]	(20,->)	-	-	-	-	-
Fracc	-	-	-	-	{0}	{0}	{0}	{0}	{0}
Party	-	-	-	-	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}
Wrkrs	-	-	-	-	{0}	{0}	{0}	{0}	{1}
Edu	-	-	-	-	(<-,2]	(2,4]	(2,4]	(4,->)	-
Gndr	-	-	-	-	-	{1}	{2}	-	-
Distant	.0897	.4217	.7568	.5510	.6750	.8736	.9792	.7953	.9259
Neighb	.9103	.5783	.2432	.4490	.3250	.1264	.0208	.2047	.0741
N	156	166	37	196	40	87	48	298	54
Rule#	1	2	3	4	5	6	7	8	9

The most important conditioning variable in this choice process is the number of days already planned for more distant destinations. When *at least one day has already been planned at a more distant destination*, the probability of selecting another more distant destination decreases dramatically (R29-34; this does not apply, however, to holidays of 9 or more days that are planned outside school holiday periods (R33)). This indicates that people are often not inclined to spend more than one holiday at more distant tourist destinations.

Table 10.3 Continued - neighbouring and more distant countries

#Drem	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{2,3}
Lngth	(2,4]	(2,4]	(2,3]	(3,4]	(2,4]	(4,->)	(4,->)	(<-,1]
Fracc	{0}	{1}	{1}	{1}	{1}	{0}	{1}	-
Party	{3,4,5,99}	-	-	-	-	-	-	-
Age	-	(<-,42]	(42,->)	(42,->)	(42,->)	-	-	-
HolP	-	-	{0}	{0}	{1}	-	-	-
Distant	.5960	.7016	.7200	.4878	.4257	.9592	.6956	.0536
Neighb	.4040	.2984	.2800	.5122	.5743	.0408	.3044	.9464
N	57	191	50	41	101	49	46	112
Rule#	10	11	12	13	14	15	16	17

When, on the other hand, *no days are planned for a more distant destination*, the probability of selecting one amounts to 54.56% (R1-28). Under these conditions, the presence of children in the household is important, followed by the duration of the holiday. Regardless of the household types, *short breaks* are usually spent in one of the neighbouring countries (R1+17). For *households without children*, medium long holidays are more likely to be spent at more distant destinations (R2-4), except when the tourist does not work (R2). Longer holidays (9+ days) are also more likely to be spent at more distant destinations (R5-16), except for long holidays by people who are older than 42 years and who own a tourist free accommodation (R13+14), and for extended holidays during school holidays (R14).

Table 10.3 Continued - neighbouring and more distant countries

#Drem	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Chld	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}
Lngth	(1,2]	(1,2]	(1,2]	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]
Fracc	-	-	-	{0}	{1}	{0}	{0}	{1}
#DT	(<-,1] + {99}	(1,->) \ {99}	(1,->) \ {99}	-	-	-	-	-
Prov	-	{1,2,4,6,8,10}	{3,5,7,9,11,12}	-	-	-	-	-
Seas+	-	-	-	{1,2,6}	{1,2,6}	{3,4,5}	{3,4,5}	{3,4,5}
HHzs	-	-	-	-	-	(<-,3]	(3,->)	-
Distant	.6000	.5111	.1579	.9167	.7391	.7143	.4167	.2973
Neighb	.4000	.4889	.8421	.0833	.2609	.2857	.5833	.7027
N	65	45	114	60	46	35	84	74
Rule#	18	19	20	21	22	23	24	25

For *households with children who have not planned any days at more distant destinations yet*, more distant destinations are also given priority for holidays of at least 5 days (67.69% vs. 32.31%; R18-28). Under these conditions, however neighbouring countries are preferred more when the holiday under consideration is a medium long holiday, the tourist resides in Drenthe, Flevoland, Utrecht, South-Holland, North-Brabant or Limburg and makes at least one day-trip per year (or has a missing value on this variable) (R20). Also, neighbouring countries are (slightly) more preferred for extended holidays during the summer season when people own a tourist accommodation without a permanent place (R25), or when people do not own such an accommodation and live in a household of at least four people (R24). Finally, neighbouring countries are also favourite for long and extra long holidays by tourist from larger households (4 or more people) who own a tourist accommodation without a permanent place (R28).

Table 10.3 Continued - neighbouring and more distant countries

#Drem	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)
Chld	{2,3}	{2,3}	{2,3}	-	-	-	-	-	-
Lngth	(3,->)	(3,->)	(3,->)	(<-,1]	(<-,1]	(1,2]	(1,2]	(2,->)	(2,->)
Fracc	-	{0}	{1}	-	-	-	-	-	-
Party	-	-	-	-	-	{1,2,5}	{3,4,99}	-	-
HolP	-	-	-	-	-	-	-	{0}	{1}
HHzs	(<-,3]	(3,->)	(3,->)	-	-	-	-	-	-
#Dnei	-	-	-	(<-,12]	(12,->)	-	-	-	-
Distant	.7188	.5281	.3729	.0189	.1017	.4969	.2237	.6296	.4462
Neighb	.2812	.4719	.6271	.9811	.8983	.5031	.7763	.3704	.5538
N	64	89	236	212	59	159	76	54	65
Rule#	26	27	28	29	30	31	32	33	34

10.4.3 The DT for the choice of the more distant destinations

The DT for the choice the more distant tourist destinations was induced using 1610 observed holidays in the CVO-data set. It comprises 33 decision rules and is presented in Table 10.4. In this table, “SpPo” indicates the probability that Spain or Portugal will be selected, “AuSw” record the popularity of Austria and Switzerland, “ItGr” denotes holidays to Italy and Greece, “EaEu” indicates holidays to (former) Eastern European countries including the (former) Yugoslavia, Soviet Union, Czechoslovakia, Hungary, Bulgaria, Rumania, and Poland, “SEM” marks the popularity of South East Mediterranean countries such as Malt, Turkey, Morocco, Cyprus and Tunisia, “ScaD” denotes holidays to Scandinavia and Denmark, and “Other”, includes all other destinations outside Europe.

Table 10.4 DT for the choice of more distant destinations

Ski	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Lngth	(<-,2]	(<-,2]	(<-,2]	(<-,2]	(<-,2]	(<-,2]	(<-,2]	(2,3]
Seas+	{1}	{2,6}	{2,6}	{2,6}	{2,6}	{3,4,5}	{3,4,5}	{1}
Prov	-	{1,4,8}	{2,5,9,10,11}	{2,5,9,10,11}	{3,6,7,12}	-	-	-
Wrkhr	-	-	(<-,36]	(36,->)	-	-	-	-
Party	-	-	-	-	-	{1,2,5,99}	{3,4}	-
SpPo	.2319	.3455	.1778	.4500	.3220	.2619	.2000	.1803
AuSw	.4348	.0000	.0444	.0500	.1864	.3333	.2000	.4590
ItGr	.0435	.2364	.1111	.0500	.1356	.2143	.0667	.0164
EaEu	.1884	.0545	.2000	.0500	.1186	.0952	.0000	.0820
SEM	.0580	.0727	.3778	.1000	.0508	.0476	.0000	.0984
ScaD	.0290	.1455	.0667	.0500	.1525	.0476	.5333	.0000
Other	.0144	.1454	.0222	.2500	.0341	.0001	.0000	.1639
N	69	55	45	20	59	42	15	61
Rule#	1	2	3	4	5	6	7	8

The *possession of skis* is important because this increases the probability of going to typical winter sports countries such as Austria and Switzerland, in particular during the *winter* (R26+27). Strikingly, for people who own skis, these countries are often also very popular during the other seasons (R28-33). However, sunny destinations like Italy, Greece, Spain and Portugal are very popular during the *spring school holidays* (R29), the *mid part of the summer* (R31), and *autumn* (R33).

Table 10.4 Continued - more distant destinations

Ski	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Lngth	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]
Seas+	{2,6}	{2,6}	{2,6}	{2,6}	{3,4,5}	{3,4,5}	{3,4,5}	{3,4,5}	{3,4,5}
Inc	(<-,3]	(3,6]	(6,7]	(7,->)	-	-	-	-	-
Age	-	-	-	-	(<-,32]	(<-,32]	(32,45]	(32,45]	(45,->)
Csize	-	-	-	-	(<-,3]	(3,->)	-	-	-
WrkW	-	-	-	-	-	-	(<-,1]	(1,->) + {9}	-
SpPo	.2647	.3645	.6364	.2128	.1818	.3700	.0588	.5455	.1290
AuSw	.0294	.0841	.0303	.1489	.0909	.0700	.3529	.0909	.3387
ItGr	.3529	.2243	.0909	.1702	.3636	.3400	.2941	.1818	.1129
EaEu	.1029	.1028	.0303	.0213	.2121	.0700	.0588	.0000	.2258
SEM	.1618	.0561	.0909	.1064	.1515	.1100	.0000	.0000	.0806
ScaD	.0294	.0187	.0000	.0000	.0000	.0200	.0588	.1818	.0806
Other	.0589	.1495	.1212	.3404	.0001	.0200	.1766	.0000	.0324
N	68	107	33	47	33	1000	17	22	62
Rule#	9	10	11	12	13	14	15	16	17

For people who *do not own skis*, the duration of the holiday is important to the choice of more distant destinations. For *extra long holidays*, the most distant destinations ("other") are often selected by young people and people who have a job (R24), whereas people without a paid job prefer destinations like Italy, Greece, Spain and Portugal (R25).

Table 10.4 Continued - more distant destinations

Ski	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Lngth	(3,4]	(3,4]	(3,4]	(3,4]	(3,4]	(3,4]	(4,->)	(4,->)
Seas+	{1,6}	{2,5}	{2,5}	{3,4}	{3,4}	{3,4}	-	-
Edu	-	(<-,2]	(2,->)	-	-	-	-	-
Sclss	-	-	-	(<-,2]	(<-,2]	(2,->)	-	-
Pair[a]	-	-	-	(<-,19.123]	(19.123,->)	-	-	-
Work	-	-	-	-	-	-	{0,1,2}	{3}
SpPo	.2273	.3333	.2979	.1944	.1875	.4054	.0789	.4474
AuSw	.0152	.4000	.0213	.2083	.0313	.1892	.0263	.0000
ItGr	.0606	.0000	.2234	.3056	.0938	.0811	.0000	.2630
EaEu	.0152	.0000	.0532	.1250	.0625	.2568	.0263	.0000
SEM	.0152	.0667	.0213	.0278	.0313	.0000	.0526	.1053
ScaD	.0000	.0000	.0851	.0833	.1875	.0135	.0526	.1316
Other	.6665	.2000	.2978	.0556	.4061	.0540	.7633	.0527
N	66	15	94	72	32	74	38	38
Rule#	18	19	20	21	22	23	24	25

[a] Rounded off at 3 decimal fractions

For shorter holidays (2-28 days) pursued by tourists without skis, the season during which the holiday is scheduled is the next most significant condition. For *short breaks, medium long holidays and extended holidays during the winter*, Austria and Switzerland are again the most popular destinations (R1+8). Apparently, these tourist rent skis on the spot. For *short breaks and medium long during spring and autumn*, Spain and Portugal are in demand (R2-5), where countries such as Malt, Turkey, Morocco, Cyprus and Tunisia are also very popular with people from Friesland, Flevoland, South-Holland, Zeeland and North-Brabant who work 36 hours per week or less (R3). For *short breaks and medium long during the summer season,* Scandinavia and Denmark are the preferred destinations for travel parties with children aged 0-5 or 15+ years and travel parties of at least 9 members (R5), whereas Austria and Switzerland, Spain and Portugal and Italy and Greece are popular destinations for the other travel parties (R6).

For *extended holidays during spring and autumn* of tourists without skis, the net household income conditions the more distant destinations choices, where

Italy and Greece and Spain and Portugal are preferred by tourists from households with incomes up to € 25,000 (R9+10), Spain and Portugal are by far the most popular with tourists from households with incomes between € 25,000 and € 29,545, and the most distant destinations ("other") are more likely to be selected by the most affluent tourists (R12). *Extended holidays during the summer*, finally, are most likely to be spent in Italy and Greece and Spain and Portugal and Austria and Switzerland (R13-17), where the latter two countries are more popular with people between 32 and 45 years old who do not work during the weekends (R15), and people older than 45 years (R17).

Table 10.4 Continued - more distant destinations

Ski	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Lngth	(<-,3]	(3,->)	-	-	-	-	-	-
Seas+	{1}	{1}	{2}	{2}	{3,4}	{3,4}	{5}	{6}
HolP	-	-	{0}	{1}	-	-	-	-
Fracc	-	-	-	-	{0}	{1}	-	-
SpPo	.0686	.0000	.2542	.1935	.1034	.1591	.0952	.2931
AuSw	.7745	.5333	.3220	.0968	.5172	.0909	.3333	.0517
ItGr	.0686	.0000	.0847	.2903	.1379	.3409	.2381	.1897
EaEu	.0392	.0000	.0169	.1613	.0345	.1364	.0000	.0517
SEM	.0392	.0000	.1186	.0968	.0345	.0682	.1429	.0517
ScaD	.0000	.0000	.0508	.0323	.1034	.1136	.0000	.0690
Other	.0099	.4667	.1528	.1290	.0691	.0909	.1905	.2931
N	102	15	59	31	29	44	21	58
Rule#	26	27	28	29	30	31	32	33

Finally, *long holidays* by people who do not own skis often go to the most distant destinations ("other"; 27.20%) or Spain and Portugal (27.76%; R18-23), in particular during the autumn and winter (R18), and during the begin and mid part of the summer by people of the two highest social classes who have an above average propensity to travel by air (> 19.1230; R22).

10.4.4 The DT for the choice of neighbouring countries

Given the choice to visit a (geographically and/or culturally) neighbouring country, Table 10.5 describes the conditions under which people decide to go to France (idem), Belgium or Luxembourg ("BLux"), Germany ("Germ") and England, Wales, Scotland or (Northern) Ireland ("UKI"). The DT comprises 40 decision rules and is based on 1656 observations.

Table 10.5 DT for the choice of the four neighbouring countries

#TGer	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Prio4	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]
Lngth	(<-,1]	(<-,1]	(1,2]	(1,2]	(1,2]	(1,2]	(2,3]	(2,3]	(2,3]	(2,3]
HolP	{0}	{1}	-	-	-	-	-	-	-	-
Age	-	-	(<-,9]	(9,40]	(9,40]	(40,->)	-	-	-	-
Gndr	-	-	-	{1}	{2}	-	-	-	-	-
Party	-	-	-	-	-	-	{1,2}	{1,2}	{3,4,5,99}	{3,4,5,99}
Fracc	-	-	-	-	-	-	{0}	{1}	-	-
Prov	-	-	-	-	-	-	-	-	{2,3,4,5,9}	{1,6,7,8,10,11,12}
France	.4667	.4194	.0556	.3158	.1471	.2063	.2619	.6364	.2745	.7297
BLux	.4000	.0645	.7778	.3684	.5882	.1905	.0714	.1818	.3529	.1622
Germ	.0667	.3226	.1111	.3158	.1765	.5238	.3333	.0682	.3726	.0811
UKI	.0666	.1935	.0555	.0000	.0882	.0794	.3334	.1136	.0000	.0270
N	15	31	18	38	34	63	42	44	51	74
Rule#	1	2	3	4	5	6	7	8	9	10

The most important condition variable in the decision which neighbouring country to visit is the number of holidays already planned for Germany. When *at least one holiday has been planned for Germany*, the next holiday is most likely to go to Germany again or to Belgium and/or Luxembourg, depending on the number of holidays already planned for the latter two countries, the possession of a tourist accommodation without a permanent place and the holiday region of the tourist (R37-40).

Table 10.5 Continued - neighbouring countries

#TGer	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Prio4	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,2]	(1,2]
Lngth	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(3,->)	(<-,2]	(<-,2]
Age	-	-	-	-	-	-	-	(<-,9]	(9,28]
Pair[a]	(<-, 13.896]	(<-, 13.896]	(<-, 13.896]	(<-, 13.896]	(13.896, 16.865]	(16.865, 21.453]	(21.453, ->)	-	-
Inc	(<-,3]	(3,4]	(4,7]	(7,->)	-	-	-	-	-
#TBLx	-	-	-	-	-	-	-	(<-,0]	(<-,0]
HHsz	-	-	-	-	-	-	-	-	(<-,3]
France	.3846	.5806	.7447	.9194	.6316	.4375	.8000	.1563	.3448
BLux	.3846	.0968	.0709	.0161	.0000	.0000	.0571	.5625	.3448
Germ	.1538	.3226	.1064	.0484	.0684	.1250	.0000	.2500	.1034
UKI	.0770	.0000	.0780	.0161	.3000	.4375	.1429	.0312	.2070
N	26	31	141	62	19	16	35	32	29
Rule#	11	12	13	14	15	16	17	18	19

[a] Rounded off at 3 decimal fractions

Table 10.5 Continued - neighbouring countries

#TGer	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Prio4	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(2,->)
Lngth	(<-,2]	(<-,2]	(<-,2]	(<-,2]	(<-,2]	(<-,2]	(2,->)	(2,->)	(2,->)	(<-,1]
HolP	-	-	-	-	-	-	{0}	{1}	{1}	-
Age	(9,28]	(9,28]	(28,52]	(52,->)	(52,->)	-	-	-	-	(<-,17]
#ABLx	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	-	-	-	(<-,0]
HHsz	(3,4]	(4,->)	-	-	-	-	-	-	-	-
Tcost	-	-	-	(<-,1050]	(1050,->)	-	-	-	-	-
#TFra	-	-	-	-	-	-	-	(<-,0]	(0,->)	-
France	.5862	.4000	.2388	.0435	.2917	.0870	.6750	.5263	.7647	.5000
BLux	.0345	.4000	.3657	.2174	.0833	.2609	.0000	.1579	.0588	.4615
Germ	.3448	.2000	.2836	.7391	.2917	.2609	.1500	.3158	.0588	.0385
UKI	.0345	.0000	.1119	.0000	.3333	.3912	.1750	.0000	.1177	.0000
N	29	15	134	23	24	23	40	38	17	26
Rule#	20	21	22	23	24	25	26	27	28	29

Alternatively, *when no holidays have been planned for Germany* yet, the priority and duration of the holiday under consideration is important. When the *most important holiday is a short break*, France is the most preferred destination, while Belgium and Luxembourg are in second place outside school holiday periods (R1) and Germany is the first runner-up during school holiday periods (R2). In contrast, when the *most important holiday is a medium long holiday*, Belgium and Luxembourg are very popular (47.22%) and Germany and France are almost equally preferred (25.00% and 23.61% respectively; where older people often prefer Germany; R3-6).

Table 10.5 Continued - neighbouring countries

#TGer	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)	(0,->)	(0,->)
Prio4	(2,->)	(2,->)	(2,->)	(2,->)	(2,->)	(2,->)	(2,->)	-	-	-	-
Lngth	(<-,1]	(<-,1]	(<-,1]	(1,2]	(2,->)	-	-	-	-	-	-
Age	(17,46]	(46,52]	(52,->)	-	-	-	-	-	-	-	-
#TBLx	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,1]	(1,->)	(<-,0]	(<-,0]	(<-,0]	(0,->)
Peracc	-	-	-	-	-	-	-	{0}	{0}	{1}	-
Region	-	-	-	-	-	-	-	{1,3}	{2}	-	-
France	.2947	.3000	.1034	.2816	.6000	.1739	.0333	.1250	.1000	.1429	.0667
BLux	.4000	.2667	.4828	.2136	.0400	.5217	.9333	.1944	.5500	.0000	.8000
Germ	.1474	.0667	.2759	.2913	.1600	.1304	.0000	.6250	.2000	.8571	.1333
UKI	.1579	.3666	.1379	.2135	.2000	.1740	.0334	.0556	.1500	.0000	.0000
N	95	30	29	103	25	46	30	72	20	21	15
Rule#	30	31	32	33	34	35	36	37	38	39	40

When the *second most important holiday is a short break or a medium long holiday, and the holiday with the highest profiling priority is planned for Belgium or Luxembourg*, the UK and Ireland are the most preferred holiday destinations (R25). When the holiday with the highest profiling priority is *not planned for Belgium or Luxembourg*, the UK and Ireland are only favoured when the tourist is at least 53 years old, and (s)he has already spent NLG 1050 or more on the previously planned holiday (R24). In all other cases, people under 53 prefer Belgium and Luxembourg and France (R18-22), while older people prefer Germany for their second most important holiday (R23).

Finally, when the *most or second most important holidays last at least 9 days*, France by far is the most important destination (R7-14+26-28). The only two exceptions to this rule are extended holidays by single travellers and/or adults only who do not own a non-permanent tourist accommodation (R7), and extended holidays by all other travel parties by a tourist who resides in Friesland, Drenthe, Overijssel, Flevoland or South-Holland (R.9).

10.4.5 The DT for the choice of Dutch metaregions

Once a tourist has decided to have a domestic holiday, (s)he has to determine the Dutch tourist area. *MERLIN* assumes the tourist to first select a metaregion. Next, the actual tourist area is considered.

Table 10.6 DT for the choice of the metaregion within the Netherlands

#DLSo	(<-,24]	(<-,24]	(<-,24]	(<-,24]	(<-,24]	(<-,24]	(<-,24]	(<-,24]
#TOth	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
#TWat	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]
Region	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
#DLNo	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Prov	{1,4,7}	{1,4,7}	{1,4,7}	{2,3}	{5,6}	{5,6}	{8,9,10,11,12}	{8,9,10,11,12}
Cst	{1,6}	{1,6}	{2,3,4,5}	-	-	-	-	{1,4}
#Tp	(<-,2]	(2,->)	-	-	-	-	-	-
Chld	-	-	-	-	{1}	{2,3}	-	-
Seas+	-	-	-	-	-	-	{1}	{2,3,4,5,6}
Water	.2963	.3333	.5588	.4046	.4554	.4211	.1207	.1862
Lnorth	.4815	.2632	.2647	.4509	.2143	.3450	.2759	.5425
Lsouth	.1852	.2632	.0732	.1214	.2500	.2222	.5690	.2267
Other	.0370	.1403	.1033	.0231	.0803	.0117	.0344	.0446
N	243	57	68	173	112	171	58	247
Rule#	1	2	3	4	5	6	7	8

Table 10.6 Continued - metaregion within the Netherlands

#DLSo	(<-,24]	(<-,24]	(<-,24]	(<-,24]	(<-,24]	(<-,24]	(<-,24]	(<-,24]
#TOth	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
#TWat	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,0]	(<-,0]	(0,1]
Region	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}
Peracc	{0}	{0}	{0}	{0}	{0}	{1}	{1}	{1}
#DLNo	(<-,0]	(<-,0]	(0,12]	(0,12]	(12,->)	(<-,12]	(<-,12]	(<-,12]
Prov	{8,9,10,11,12}	{8,9,10,11,12}	-	-	-	-	-	-
Cst	{2,6}	{3,5}	-	-	-	-	-	-
Seas+	{2,3,4,5,6}	{2,3,4,5,6}	-	-	-	-	-	-
Wrkhr	-	-	(<-,24]	(24,->)	-	-	-	-
Party	-	-	-	-	-	{1,2,3,4,5}	{99}	-
Water	.3245	.3333	.1549	.1782	.1333	.3438	.2045	.5778
Lnorth	.4176	.2549	.6127	.3960	.7250	.2917	.6477	.3556
Lsouth	.2473	.3137	.1761	.3960	.0917	.2604	.1364	.0444
Other	.0106	.0981	.0563	.0298	.0500	.1041	.0114	.0222
N	182	51	142	101	120	96	88	45
Rule#	9	10	11	12	13	14	15	16

MERCIN distinguishes four metaregions within the Netherlands: "water", "land north" and "land south", and "other". Based on the 3855 observed domestic holidays, Table 10.6 displays the 32 decision rules that describe the choices between these regions.

The three most important factors that condition the choice of metaregion are:

- the number of days already planned for the metaregion "land south";
- the number of holiday already planned for the metaregion "other"; and
- the number of holiday already planned for the metaregion "water".

One by one these variables indicate the importance of destination loyalty. Under certain conditions, however, variety-seeking behaviour is also evident. When *at least 25 days have already been planned for the Metaregion "land south",* for instance, the tourist is almost certain to return to this metaregion for his next holiday(s) (R32). When *24 days or less days have been planned for the Metaregion "land south"*, and the tourist has already planned *one or more holiday to the metaregion "other"*, the tourist is likely to return to this metaregion when travelling alone, with "other" children (0-5 years or 15+ years) or with unknown travel companions (R30; destination loyalty). However, when the tourist is travelling with adults only, with schoolchildren or with parties of at least 9 members, the preferences for the metaregions are more diffuse, where "land north" is the most popular destination (R31). Finally, when *24 days or less days have been planned for the metaregion "land* *south"*, the tourist has *not yet planned any holiday to the metaregion "other"*, and (s)he *has already planned 2 or more holidays for the metaregion "water"*, the tourist is most likely to return to the latter metaregion (R28+29), again indicating destination loyal behaviour.

Table 10.6 Continued - metaregion within the Netherlands

#DLSo	(<-,24]	(<-,24]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
#TOth	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
#TWat	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]
Region	{1,2}	{1,2}	{3}	{3}	{3}	{3}	{3}
Peracc	{1}	{1}	-	-	-	-	-
#DLNo	(12,->)	(12,->)	-	-	-	-	-
Prov	-	-	{1,2,3,4,5,6, 7,8,11,12}	{1,2,3,4,5,6, 7,8,11,12}	{1,2,3,4,5,6, 7,8,11,12}	{1,2,3,4,5,6, 7,8,11,12}	{1,2,3,4,5,6, 7,8,11,12}
Seas+	-	-	{1,2,3}	{4,5,6}	-	-	-
Party	-	-	{1,3,99}	{1,3,99}	{2}	{2}	{2}
Fracc	{0}	{1}	-	-	-	-	-
Prio4	-	-	-	-	(<-,1]	(1,2]	(2,->)
Water	.0000	.0147	.3284	.6481	.4496	.4909	.3125
Lnorth	.9920	.9706	.2687	.1296	.4082	.1091	.4375
Lsouth	.0080	.0000	.3284	.2037	.1224	.3091	.1458
Other	.0000	.0147	.0745	.0186	.0198	.0909	.1042
N	249	68	67	108	49	55	48
Rule#	17	18	19	20	21	22	23

When *24 days or less days have been planned for the metaregion "land south"*, *no holiday have been planned to the metaregion "other" yet*, and *at most one holiday has been planned for the metaregion "water"*, there are still 27 decision rules left to determine the preferences for the four Dutch metaregions (R1-27). Under these conditions, the holiday region of the tourist is the next most significant condition variable. When at least one day has been planned for "land south", tourists from the *southern holiday region* of the Netherlands, are most likely to visit this region again (R26+27). If, in contrast, no days have been planned for "land south", "land north" is the most popular choice for these tourists (R19-25).

Under the above-mentioned conditions, the decision by tourists from the *northern and mid holiday regions* is affected most by the *possession of a permanent tourist accommodation*. When these tourists own such an accommodation, "land north" is the preferred holiday destination (R14-19), except when the tourist has spent less than 13 days in that metaregion, and (s)he is travelling with a known travel party (R14), or (s)he has planned at least one holiday in the metaregion "water" (R16). In the latter two cases, "water" is the most popular option.

When tourists from the northern and mid holiday region of the Netherlands *do not own a tourist accommodation with a permanent place*, the number of days already planned for “land north” is important, where people who have planned at least one day in this metaregion, are inclined to return to this region (R11-13). For people who have not planned any days for “land north” yet, the province of residence is important. Under these conditions, people from Groningen, Overijssel and Utrecht often prefer the metaregion “water”(R1-3), except when their civil state is married or unknown, and they have only two (or less) holiday left to profile (R1). Tourists from Friesland and Drenthe prefer the metaregion “land north”, although “water” is a good second (R4). Inhabitants of Flevoland and Gelderland prefer the metaregion “water” (R5+6). And people from the other provinces, finally, mostly decide to visit “land south” during the winter (R7), whereas “land north” is more popular during the other seasons (R7-10) (except for people who are divorced and/or single (R10)).

Table 10.6 Continued - metaregion within the Netherlands

#DLSo	(<-,0]	(<-,0]	(0,24]	(0,24]	(<-,24]	(<-,24]	(<-,24]	(<-,24]	(24,->)
#TOth	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)	-
#TWat	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,->)	(1,->)	-	-	-
Region	{3}	{3}	{3}	{3}	-	-	-	-	-
Peracc	-	-	-	-	-	-	-	-	-
#DLNo	-	-	-	-	-	-	-	-	-
Prov	{1,2,3,4,5,6, 7,8,11,12}	{9,10}	-	-	-	-	-	-	-
Cst	-	-	-	-	-	-	-	-	-
#Tp	-	-	(<-,1]	(1,->)	-	-	-	-	-
Party	{4,5}	-	-	-	{1,99}	{2,3,4,5}	{1,4,99}	{2,3,5}	-
Water	.3677	.2614	.2639	.1034	.9864	.4407	.2444	.1475	.0000
Lnorth	.1677	.4216	.3056	.0862	.0045	.3390	.0778	.3607	.0085
Lsouth	.4581	.2484	.3889	.8103	.0045	.1525	.0111	.2623	.9915
Other	.0065	.0686	.0416	.0001	.0046	.0678	.6667	.2295	.0000
N	155	306	72	58	220	59	90	61	236
Rule#	24	25	26	27	28	29	30	31	32

10.4.6 The DT for the choice of tourist areas within the Dutch metaregion “Land south”

Based on the 976 observed holidays to the metaregion “land south“, Table 10.7 represents the 14 decision rules that determine whether people will visit the tourist area “land south-west” (LSWest), which includes the CBS-areas West & Mid Brabant, or the tourist area “land south-south” (LSSouth), which includes the CBS-areas East Brabant, Nijmegen and (north, mid and south) Limburg.

Table 10.7 DT for the choice of tourist area within the metaregion “Land south”

Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}
#TLSW	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
HolP	{0}	{0}	{0}	{0}	{0}	{0}	{1}
Wrkrs	{0}	{0}	{0}	{0}	{0}	{1}	-
#TLSS	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	-	-
Pcar[a]	(<-,55.898]	(55.898,65.184]	(65.184,85.596]	(85.596,->)	-	-	-
LSWest	.2000	.5000	.2386	.5490	.0645	.5588	.2428
LSSouth	.8000	.5000	.7614	.4510	.9355	.4412	.7572
N	35	34	88	51	31	34	346
Rule#	1	2	3	4	5	6	7

[a] Rounded off at 3 decimal fractions

The most significant condition in the choice between the tourist areas of “land south” is the *possession of a tourist accommodation with a permanent place*. When people own such an accommodation, the majority of holidays to “land south” are spent in the western part of this metaregion (R9-14). This is probably due to the fact that this part of the Netherlands offers many facilities for permanent tourist accommodations. Exceptions occur, however, for people from the northern holiday region (R9), and for people from the other regions who have already planned at least one holiday in a neighbouring country (R11).

When people do *not own a tourist accommodation with a permanent place*, the *number of holidays already planned for “land south-west”* is important, where, again, destination loyalty is evident when people have visited this tourist area before (R8). If, on the other hand, *no holidays have been planned for this tourist area*, “land south-south” is the preferred tourist destination (R1-7), in particular during school holiday periods (R7), or outside these periods when people are not restricted to certain holiday seasons, and at least one holiday has already been planned to “land south-south” (R5) or when people have a rather low (smaller than 85.6%) propensity to travel by car (R1-3).

Table 10.7 Continued - tourist area within the metaregion “Land south”

Peracc	{0}	{1}	{1}	{1}	{1}	{1}	{1}
#TLSW	(0,->)	-	(<-,0]	(<-,0]	(0,4]	(4,->)	(4,->)
Region	-	{1}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}
#Tnei	-	-	(<-,0]	(0,->)	-	-	-
#TLSo	-	-	-	-	-	(<-,6]	(6,->)
LSWest	.6512	.2222	.7576	.1304	.9194	.9565	1.000
LSSouth	.3488	.7778	.2424	.8696	.0806	.0435	.0000
N	43	27	33	23	63	23	146
Rule#	8	9	10	11	12	13	14

10.4.7 The DT for the choice of tourist areas within the Dutch metaregion "Land north"

Given the choice to visit the metaregion "land north", the next destination choice set comprises the choice options "land north-mid" (LNMid), including the CBS-areas Utrecht, 't Gooi and the Veluwe, "land north-east" (LNEast), including the areas Achterhoek, Twente, Salland and the Vecht-area, and "land north-north" (LNNorth), which covers the sand areas in the North of the Netherlands. Using the 1490 observed holidays in this metaregion, Table 10.8 shows the 14 induced decision rules that describe the preferences for these areas under various conditions. The three most influencing condition variables recapitulate the numbers of holidays already planned for the three tourist areas within the metaregion "land north". When at least one holiday has been planned for the northern part of the metaregion "land north", subsequent domestic holidays too will often go to this tourist area (R12-14). When *no holiday has been planned for "land north-north"*, but *at least one holiday has been planned for "land north-mid"*, holidays with lower profiling priorities will often also go the latter tourist area (R8-11). Finally, when *no holiday has been planned for "land north-north" or "land north-mid"*, and *at least one holiday has been planned for the Eastern parts of the metaregion "land north"*, the destination for subsequent holidays is often "land north-east" (R6+7). Again, these decision rules indicate strong preferences for repeat visits to tourist areas within a one year period.

Table 10.8 DT for the choice of tourist area within the metaregion "Land north"

#TLNN	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
#TLNM	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
#TLNE	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)
Provin	{1,3,8,12}	{2}	{4,5,6,7}	{4,5,6,7}	{9,10,11}	-	-
Peracc	-	-	{0}	{1}	-	{0}	{1}
LNMid	.3498	.1765	.2594	.4231	.5108	.2000	.0417
LNEast	.3270	.2353	.4292	.5000	.2985	.5500	.9479
LNNorth	.3232	.5882	.3114	.0769	.1907	.2500	.0104
N	263	51	212	26	325	60	96
Rule#	1	2	3	4	5	6	7

When *no tourist trips have been planned for the metaregion "land north"*, destination choices within this region are strongly conditioned by the *province of residence* of the tourist. Tourists from Groningen, Drenthe, North-Holland and Limburg equally prefer the three parts of the metaregion (R1). Tourists from Friesland, on the other hand, have stronger preferences for the northern part of “land north” (R51). People from Overijssel, Flevoland, Gelderland and Utrecht preferably go to “land north-east”, but the mid and northern parts are also popular (R3+4). Tourists from South-Holland, Zeeland and North-Brabant, finally, prefer the mid part, while the eastern part is the first runner-up (R5). Generally, rules comprising the province of residence (R1-5) indicate that, given the choice for “land north”, tourists often select the tourist area that is closest to their home.

Table 10.8 Continued - tourist area within the metaregion “Land north”

#TLNN	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)	(0,->)
#TLNM	(0,->)	(0,->)	(0,->)	(0,->)	-	-	-
Partu	{1,99}	{1,99}	{2,3}	{4,5}	{1,2,3,4,5}	{99}	{99}
Fracc	{0}	{1}	-	-	-	{0}	{1}
LNMid	1.000	.8649	.5738	.5122	.2987	.0000	.0882
LNEast	.0000	.1081	.3607	.3195	.3377	.0000	.0294
LNNorth	.0000	.0270	.0655	.1683	.3636	1.000	.8824
N	117	37	61	41	77	91	34
Rule#	8	9	10	11	12	13	14

10.4.8 The DT for the choice of tourist areas within the Dutch metaregion “Water”

The final DT for destination choices describes the choice between the three tourist areas within the Dutch metaregion “water”. The choice set comprises “water-north” (WNorth), i.e. the water areas (beaches, lakes) in the North, “water-mid” (WMid), which includes the water areas in the middle of the Netherlands, and “water-north sea” (WNsea), which comprises the North Sea coastal area.

Table 10.9 DT for the choice of tourist area within the metaregion “Water”

#TWM	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
#TWNs	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Region	{1}	{1}	{1}	{1}	{2,3}	{2,3}	{2,3}
#TWN	(<-,0]	(<-,0]	(0,->)	(0,->)	-	-	-
Prov	{1,2,3,5}	{4,6,7,8,9,10,11,12}	-	-	-	-	-
Party	-	-	{2,3,5}	{1,4,99}	-	-	-
Pair[a]	-	-	-	-	(<-,13.293]	(<-,13.293]	(<-,13.293]
#Trem	-	-	-	-	(<-,0]	(<-,0]	(0,->)
Peracc	-	-	-	-	{0}	{1}	-
WNorth	.7355	.4935	.7879	1.000	.2338	.1026	.4318
WMid	.0413	.1234	.0606	.0000	.1602	.4359	.0909
WNsea	.2232	.3831	.1515	.0000	.6060	.4615	.4773
N	121	154	33	59	231	39	88
Rule#	1	2	3	4	5	6	7

[a] Rounded off at 3 decimal fractions

Table 10.9 shows the set of 15 exclusive and exhaustive decision rules that describe the choice processes for these tourist areas. When *at least one holiday has already been planned for the "water-mid"-area*, subsequent holidays are likely to be spent in this area too (R14+15). Similarly, when no holidays have been planned for "water-mid", and *at least one holiday has already been planned for "water-north sea"*, holidays with lower profiling priorities will probably also go to the "water-north sea"-area (R11-13).

Table 10.9 Continued - tourist area within the metaregion "Water"

#TWM	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)
#TWNs	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)	(0,->)	-	-
Region	{2}	{3}	{2,3}	-	-	-	-	-
Pair[a]	(13.293,->)	(13.293,->)	(13,293,->)	-	-	-	-	-
Peracc	-	-	-	{0}	{1}	{1}	-	-
Fracc	{0}	{0}	{1}	-	-	-	-	-
#TD	-	-	-	-	(<-,3]	(3,->)	-	-
Gndr	-	-	-	-	-	-	{1}	{2}
WNorth	.6000	.3944	.4783	.4203	.1000	.0000	.1463	.0270
WMid	.0909	.0563	.2391	.0435	.0250	.0000	.7317	.9459
WNsea	.3091	.5493	.2826	.5362	.8750	1.000	.1220	.0271
N	55	71	46	69	40	93	41	37
Rule#	8	9	10	11	12	13	14	15

[a] Rounded off at 3 decimal fractions

When *no holidays have yet been planned for "water-mid" or "water-north sea",* the holiday region is important in deciding which water area to visit. Under these conditions, people from the *northern holiday region* tend to spend their holiday in the northern water areas (R1-4). In contrast, people from the *mid and southern holiday regions* in the Netherlands often prefer the North Sea coastal areas (R5-10). The northern water areas are preferred, however, when people have a more than average propensity to travel by aeroplane (although this obviously does not apply to the domestic holidays under consideration), and either own a tourist accommodation without a permanent location (R10), or live in the mid holiday region and do not own such an accommodation (R8).

10.5 Statistics and validation

This final section discusses the most important statistics and the predictive abilities of the destination DTs that were discussed in the previous section. The confusion matrix and statistics for the DT for the choice of domestic or foreign holidays are presented in Table 10.10. The probabilistic DT correctly classifies 69.18% of the observations, an improvement of 37.41% compared to the probabilistic single rule model. For the individual categories, the prediction improves by 32.12% for domestic holidays (from 54.14% to 71.53%), and 44.79% for holidays abroad (from 45.86% to 66.40%). The reasonable performance of the DT is confirmed by the contingency coefficient that amounts to 0.5235 out of a maximum of 0.7071 for a 2×74 DT.

Table 10.10 Confusion matrix & statistics for DT *domestic or abroad*

	NL'	Abroad'	Total'
NL	**.7153**	.2847	.5414
Abroad	.3360	**.6640**	.4586
Total	.5414	.4586	**.6918**
N° of observations	7121		
Stopping criteria	80 before/ 35 after		
α	0.05		
N° of columns (= rules)	74		
Theta (1 col)	3855 (.5414)[a]	3585 (.5034)[b]	
Theta (74 cols)	5476 (.7690)[a]	4926 (.6918)[b]	
χ^2_{DT}(df) / Cont. coef.	2690.03 (73) / .5235		

[a] Deterministic [b] Probabilistic

Table 10.11 presents the statistics and indicators of the predictive validity of the DT for the choice between neighbouring and more distant countries. Based on 3266 observed holidays, the 34 induced choice rules are able to correctly predict 2214 observations, 481 more than the single rule-model using a probabilistic assignment rule (29.46% improvement).

Table 10.11 Confusion matrix & statistics for the DT *neighbouring or more distant countries*

	Neighbouring'	More distant'	Total'
Neighbouring	**.6523**	.3477	.5070
More distant	.3577	**.6423**	.4930
Total	.5070	.4930	**.6473**
N° of observations	3266		
Stopping criteria	80 before/ 35 after		
α	0.05		
N° of columns (= rules)	34		
Theta (1 col)	1656 (.5507)[a]	1633 (.5000)[b]	
Theta (34 cols)	2363 (.7235)[a]	2114 (.6473)[b]	
χ^2_{DT}(df) / Cont. coef.	962.08 (33) / .4770		

[a] Deterministic [b] Probabilistic

This improvement is less powerful compared to the previous DT, which is confirmed by the lower contingency coefficient (0.4770 where the maximum again equals 0.7071). However, in case of the DT for neighbouring and more distant countries, the increased predictive abilities are equally distributed over the two choice alternatives because the expected number of correctly predicted observations increases from 50.70% to 65.23% for neighbouring countries, and from 49.30% to 64.23% for more distant countries.

Table 10.12 Confusion matrix & statistics for the DT *more distant tourist destinations*

	SpPo’	AuSw’	ItGr’	EaEu’	SEM’	ScaD’	Other’	Total’
SpP0	**.3149**	.1431	.1812	.0896	.0717	.0592	.1402	.2590
AuSw	.1832	**.4222**	.1167	.0906	.0602	.0399	.0873	.2000
ItGr	.2629	.1435	**.2372**	.0973	.0787	.0591	.1214	.1634
EaEu	.2497	.2048	.1706	**.1498**	.0815	.0524	.0913	.0876
SEM	.2475	.1668	.1833	.1027	**.1342**	.0492	.1164	.0720
ScaD	.2669	.1387	.1824	.0809	.0624	**.1393**	.1293	.0578
Other	.2427	.1092	.1356	.0504	.0564	.0512	**.3545**	.1770
Total	.2590	.2000	.1634	.0876	.0720	.0578	.1770	**.2925**

N° of observations	1610	
Stopping criteria	35 before/ 15 after	
a	0.05	
N° of columns (= rules)	33	
Theta (1 col)	417 (.2590)[a]	283 (.1758)[b]
Theta (33 cols)	684 (.4248)[a]	471 (.2925)[b]
χ^2_{DT}(df) / Cont. coef.	1787.66 (192) / .6425	

[a] Deterministic [b] Probabilistic

The confusion matrix and statistics for the DT that describes the choice processes for the seven more distant tourist destinations are provided in Table 10.12. The expected number of correctly predicted observations using the 33 probabilistic decision rules amounts to 471 (29.25%).

This number may seem disappointing, but it is still a very important improvement in comparison with the probabilistic null-model, that only has an expected number of 283 (17.58%). In other words, the DT improves the number of hits by 66.43%. A comparison between the correctly predicted number of cases for each category by the decision model (diagonal of the confusion matrix) and that of a probabilistic single-rule model (utmost right column) shows that 4 out of the 7 groups of countries double the number of hits. These positive results are confirmed by the contingency coefficient that amounts to 0.6425 on a maximum of 0.9252 for this 7×34 DT.

Table 10.13 Confusion matrix & statistics for the DT *neighbouring countries*

	France'	BLux'	Germany'	UKI'	Total'
France	**.5502**	.1710	.1737	.1051	.3961
BLux	.2652	**.4178**	.2155	.1015	.2554
Germany	.2998	.2325	**.3614**	.1062	.2367
UKI	.3518	.2321	.2138	**.2023**	.1117
Total	.3961	.2554	.2367	.1117	**.4330**

N° of observations	1656	
Stopping criteria	35 before/ 15 after	
α	0.05	
N° of columns (= rules)	40	
Theta (1 col)	656 (.3961)[a]	481 (.2905)[b]
Theta (40 cols)	908 (.5481)[a]	717 (.4330)[b]
χ^2_{DT}(df) / Cont. coef.	893.17 (117) / .5919	

[a] Deterministic [b] Probabilistic

Based on 1656 observations, the 40 probabilistic decision rules for the choice of one of the neighbouring countries are able to correctly predict 717 of these observations, an improvement of 49.06% compared to the null-model (see Table 10.13). With regard to the destinations, the improvement is most significant for the UK and Ireland, because for this category, the expected number of correct predictions almost doubles (81.11%) compared to the null-model. For the other destinations the improvements are also significant, respectively 38.90% (France), 63.59% (Belgium and Luxembourg) and 52.68% (Germany). Overall, there is a reasonably strong relation between the conditions and response distributions in this DT. The contingency coefficient amounts to 0.5219 on a maximum of 0.8660 for this 4$\times$40 DT.

Table 10.14 Confusion matrix & statistics for the DT *Dutch metaregions*

	Water'	L. North'	L. South'	Other'	Total'
Water	**.4791**	.2842	.1851	.0516	.3053
L. North	.2243	**.5606**	.1733	.0419	.3568
L. South	.2232	.2647	**.4759**	.0362	.2532
Other	.2879	.2959	.1674	**.2487**	.0547
Total	.3053	. 3582	.2532	. 0547	**.4973**

N° of observations	3855	
Stopping criteria	80 before/ 35 after	
α	0.05	
N° of columns (= rules)	32	
Theta (1 col)	1491 (.3568)[a]	1195 (.3100)[b]
Theta (32 cols)	2274 (.5899)[a]	1917 (.4973)[b]
χ^2_{DT}(df) / Cont. coef.	2976.79 (99) / .6601	

[a] Deterministic [b] Probabilistic

Given the decision to stay in the Netherlands, the DT for the choice between the metaregions described the first stage of the destination choice process for domestic holidays. The confusion matrix and statistics for this 4×32 DT are presented in Table 10.14. Based on the contingency coefficient that amounts to 0.6601 out of a maximum of 0.8660, it can be concluded that there is a rather strong relationship between the conditions and the responses. Compared to the null-model, the 32 probabilistic decision rules increase the expected number of correctly predicted observations by 60.42% (from 1195 to 1917). This improvement is particularly strong for the holidays to the metaregion "other", where the number of correctly predicted observations increases from 5.47% to 24.87%.

Table 10.15 Confusion matrix & statistics for the DT *tourist areas within the Dutch metaregion "land south"*

	LSWest'	LSSouth'	Total'
LSWest	**.6871**	.3129	.4764
LSSouth	.2842	**.7158**	.5236
Total	.4764	.5236	**.7018**

N° of observations	976	
Stopping criteria	50 before/ 20 after	
α	0.05	
N° of columns (= rules)	14	
Theta (1 col)	511 (.5236)[a]	489 (.5011)[b]
Theta (14 cols)	769 (.7879)[a]	685 (.7018)[b]
χ^2_{DT}(df) / Cont. coef.	394.53 (13) / .5363	

[a] Deterministic [b] Probabilistic

Table 10.15 shows the confusion matrix and statistics for the DT for the two tourist areas of the metaregion "land south". The contingency coefficient of this 2×14 DT amounts to 0.5363 on a maximum of 0.7071, indicating a relatively strong relationship between the condition and the decision variables. The DT adds 196 correctly predicted observations to that of the probabilistic null-model, an increase of 40.08%. The improved predictive power is slightly larger for the tourist area "land south-west", but differences are small (44.23% vs. 36.71% improvement).

The confusion matrix and statistics for the DT for the tourist areas of "land north" are presented in Table 10.16. The probabilistic DT correctly classifies 51.24% of the 1491 observed holidays, an increase of 159 observations compared to the null-model (+31.49%). The improvement is particularly large for "land north-mid" and "-east" since the expected number of correctly classified observations increase by 71.04% and 65.94% respectively. Also the contingency coefficient is fairly reasonable at 0.5879 out of a maximum of 0.8165 for this 3×14 DT.

Table 10.16 Confusion matrix & statistics for the DT *tourist areas within the Dutch metaregion "land north"*

	LNNorth'	LNMid'	LNEast'	Total
LNNorth	**.4958**	.2524	.2519	.3890
LNMid	.1817	**.5564**	.2619	.3253
LNEast	.2171	.3089	**.4741**	.2857
Total	.3890	.3253	.2857	**.5124**

N° of observations	1491	
Stopping criteria	60 before/25 after	
α	0.05	
N° of columns (= rules)	14	
Theta (1 col)	580 (.3890)[a]	505 (.3388)[b]
Theta (14 cols)	870 (.5835)[a]	764 (.5124)[b]
χ^2_{DT}(df) / Cont. coef.	787.43 (26) / .5879	

[a] Deterministic [b] Probabilistic

The confusion matrix and statistics of the final destination DT are presented in Table 10.17. The 15 decision rules increase the expected number of correctly predicted observations from 453 to 659 (45.47% improvement). The improvement is particularly strong for "water-mid", where the expected number of hits increases by 182.7%. Although not as extreme, the predictive power of the DT for the other areas is also reasonable (from 32.34% for "water-North Sea" to 42.74% for "water-north"). Again, these observations are confirmed by the contingency coefficient that amounts to 0.6095 out of a maximum of 0.8165 for this 3×15 DT.

Table 10.17 Confusion matrix & statistics for the DT *tourist areas within the Dutch metaregion "water"*

	WNorth'	WMid'	WNsea'	Total
WNorth	**.5688**	.0959	.3353	.3985
WMid	.2543	**.4252**	.3205	.1504
WNsea	.2961	.1068	**.5970**	.4511
Total	.3985	.1504	.4511	**.5599**

N° of observations	1177	
Stopping criteria	60 before/25 after	
α	0.05	
N° of columns (= rules)	15	
Theta (1 col)	531 (.4511)[a]	453 (.3849)[b]
Theta (15 cols)	774 (.6576)[a]	659 (.5599)[b]
χ^2_{DT}(df) / Cont. coef.	695.60 (28) / .6095	

[a] Deterministic [b] Probabilistic

10.6 Conclusion and discussion

This chapter discussed the DTs for tourist destination choices. In essence, MERCIN distinguishes several stages in the destination choice process, the first of which is the decision to go abroad or to stay in the Netherlands. In this choice, the possession of a tourist accommodation with a permanent location and the presence of children in the household proved to be the most significant condition variables. Given the decision to go abroad, MERCIN assumes the tourist to select particular (groups of) countries following the decision to visit either neighbouring or more distant countries. Similarly, for domestic holidays, the selection of tourist areas is also a 2-staged process.

Transport mode considerations were assumed to be imporant to tourist destination choices. More specifically, the tourists' individual propensities to go on either car- or air-based holidays or to select another transport mode were entered as conditions into the CHAID-based algorithms. These propensities were calculated using a MNL-model that related the observed mode preferences (in the 1998 CVO-data) to the aggregate marginal distribution of these transport mode alternatives, the tourist's annual number of holidays and several personal and household conditions. Using this model, MERCIN can evaluate the effects of changes in the aggregate marginal distribution of these mode alternatives on tourist choices. The three "mode-propensities" were selected as condition variables in 5 of the 8 destination DTs. In the choice of more distant destinations, for instance, it was shown that people with higher propensities to travel by aeroplane, were more likely to select the most distant destinations (R21+22). However, the mode-propensity conditions were never among the most important conditions, suggesting that shifts in transport mode preferences may not have significant effects on destination choices. This will be demonstrated in chapter 12.

With regard to the other circumstances that condition destination choices, the discussion of the DTs revealed that when people have already planned one or more holidays at more distant destinations, the probability of visiting another more distant destination decreases dramatically. In contrast, in the choice between domestic metaregions and tourist areas (and to some extent also in the choices between the neighbouring countries), repeat visitation patterns were often observed. In these cases, when one or more holidays had already been planned for a particular area, the probability of visiting that destination again often increased. Apparently, within the period of one year, distant holiday destinations are one-of-a-kind experiences, whereas destinations closer to home may be selected again for holidays with lower profiling priorities.

11 Accommodation, Transport Mode and Expenditure Choices

This last empirical chapter discusses the final three tourist decisions, including the choice of accommodation, the choice of transport mode and the expenditures for each holiday in the annual tourist trip program. The first two choices are categorical choices and will be described using decision rules that are induced from observed choices by a CHAID-based algorithm. In contrast, the expenditures for each holiday are non-negative continuous choices. Following a categorisation into ten segments, however, these choices too are described using probabilistic decision rules.

11.1 Introduction

This chapter concludes the series of chapters that discuss the empirical input into the *MERLIN*-system. Given the choices of travel party, "during school holiday period or not", season and destination, several holiday decisions remain, including the choices of accommodation, transport mode, and expenditures. The first two choices are both categorical decision variables. These tourist choices are therefore induced and represented as before. In contrast, the choice of expenditures is a non-negative continuous allocation process. Various modelling approaches can be adopted to describe this choice process, including left-censored regression analysis (i.e., Tobit regression), the induction of decision rules for continuous action variables and the induction of decision rules that predict the selection of categorised ranges of expenditures. Based on both empirical and practical grounds, expenditure choices in the *MERLIN*-system are described using categorised ranges.

The next sections discuss accommodation, transport mode and expenditure choices. In contrast to the previous chapters, however, the statistics and validation of each DT are discussed directly following the discussion of each DT. Finally, summarising the major findings with regard to these tourist choices will conclude this last empirical chapter.

11.2 The question of where to stay

11.2.1 The choice set

The accommodation alternatives considered by the *MERLIN*-system are mainly based on the ownership of the facility. This dimension was selected because holidays are often cheaper and more flexible with regard to timing and location when a tourist decides to use an accommodation that is owned by him- or her-self.

Also, interdependencies between destination and transport mode choices and the ownership of the accommodation are likely because the convey of the personal accommodation may be difficult, expensive or even impossible. Third and finally, the ownership of the desired accommodation type is important to the tourism industry because they have to attract and accommodate tourists in order to stay in business. With regard to accommodations that are not owned by the tourist, *MERCIN* distinguishes between “hotels, motels, pensions, apartments, and rooms without pension” and “other accommodations not owned by the tourist”. For ease of readability, the first category will be referred to as “hotels” and the second as “other”. The most important accommodations in the latter category (comprising 89% of the observations) include summer cottages, holiday bungalows, first and second houses of private persons (other than the tourist) and (rented or exchanged) tents and caravans.

With regard to the accommodations that are owned by the tourist, the *MERCIN*-system distinguishes between accommodations with and accommodations without a permanent location. This distingtion is motivated by the relationship between destination and transport mode choices. Basically “permanent personal accommodations”, including caravans, tents and boats with season or permanent places and summer cottages, second houses and allotment-facilities, can often not be moved freely to other destinations. In contrast, “non-permanent personal accommodations” such as tents, caravans, boats and camper vans without season or permanent places are by definition more flexible. However, the latter accommodation type may impose restrictions on the potential transport mode.

11.2.2 The condition variables

As before, the condition variables common to all choice facets (Table A1.1 in Appendix 1) are entered into the CHAID-based algorithm for tourist accommodation choices. Also, the decisions that have been taken previously condition the tourist choices under consideration (Table A1.2). These earlier decisions include the travel party, the decision whether or not to travel during a school holiday period, the season, three conditions related to the profiling priority of the holiday, and the destination. In addition, there are nine condition variables specific to accommodation choices (Table A1.16). First, as before, the *MERCIN*-system assumes the total amount of expenditures that have already been planned for holidays with higher profiling priorities to condition subsequent accommodation choices. Second, eight summary variables represent the accommodation choices on holidays with higher profiling priorities both in terms of the number of holidays that are associated with each of the four accommodation alternatives, as well as the number of days that are involved with these decisions. These variables are able to capture patterns of recurrent and/or alternating accommodation decisions when they are present in the sample of observations.

11.2.3 The stopping criteria of the CHAID-based algorithm

As before, the α-level for predictor eligibility is set at 5%. With regard to the minimum number of observations before/after splitting the sample, the performance of various models is reported in Table A2.14 (Appendix 2). Based on these analysis, the stopping criteria for the choice of accommodation are set at 80 observations before and 35 after division because this model equals or even exceeds the aggregate performance of the models with more decision rules while the cross-validation measure is not noticeably worse. Stricter stopping criteria importantly decrease both the aggregate and cross-validation performance.

11.2.4 The generated decision table

Given the decision and condition variables and algorithm settings discussed in the previous sections, Table 11.1 presents the resulting set of 90 decision rules. In this table "hotel" refers to "hotels, motels, pensions and apartments" that are not owned by the tourist, "Nperm" denotes the probability of selecting a personal accommodation without a permanent location, "Perm" indicates the "permanent personal accommodations", and "other", finally represents all other tourist accommodations that are not owned by the tourist.

As expected the possession of tourist accommodation with a permanent location plays an important role in the choice of accommodation for holidays. *When at least one holiday* has been spent in such an accommodation, the permanent personal accommodation is by far the most popular choice (R82-90)[34]. However, this does not apply to foreign holidays with children, adults only or parties of at least 9 members (R89+90).

[34] Strikingly, rule 85 says that under these conditions people who do not own a tourist accommodation with a permanent location will still select this alternative for 88.64% of their domestic holidays with unknown travel parties. Similarly, rules 4-6, 34, 44, 48, 51, 53-54 and 64 also state that people without such an accommodation have a small chance of selecting it anyway. The only reasonable explanation for these observations is that the information on the possession of tourist commodities was collected in March during the first quarterly CVO-measurement, and that people have purchased the accommodation (or the season place) after that.

Table 11.1 DT for the choice of accommodation

#Tper	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Fracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Chld	{1}	{1}	{1}	{1}	{1}	{1}	{2,3}	{2,3}	{2,3}	{2,3}
Lngth	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,2]	(2,->)	(<-,1]	(<-,1]	(1,->)	(1,->)
Party	{1,2,99}	{1,2,99}	{1,2,99}	{3,4,5}	-	-	{1,3,4}	{1,3,4}	{1,3,4}	{1,3,4}
Wrkhr	(<-,0]	(0,36]	(36,->)	-	-	-	-	-	-	-
HolP	-	-	-	-	-	-	{0}	{1}	-	-
Car	-	-	-	-	-	-	-	-	{0}	{1}
Hotel	.7935	.5294	.7361	.3293	.2971	.1310	.1224	.2500	.0169	.0641
Nperm	.0000	.0588	.0139	.0122	.0109	.0119	.0000	.1053	.0678	.0192
Perm	.0000	.0000	.0000	.0036	.0036	.0119	.0000	.0000	.0000	.0000
Other	.2065	.4118	.2500	.6549	.6884	.8452	.8776	.6447	.9153	.9167
N	92	85	72	82	276	84	49	76	59	312
Rule#	1	2	3	4	5	6	7	8	9	10

For domestic holidays, the personal permanent accommodation is not the most popular alternative only when the travel party consists of at least 9 members and/or includes adults only or schoolchildren, and the holidays is scheduled during the autumn and winter seasons, or the first part of the summer holidays (R83).

Table 11.1 Continued - Accommodation

#Tper	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Fracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
D_NLA	{1}	{1}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Chld	{2,3}	{2,3}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Lngth	-	-	(<-,2]	(2,->)	-	-	-	-	-	-	-
Party	{2,99}	{5}	-	-	-	-	-	-	-	-	-
HolP	-	-	-	-	-	-	-	-	-	{0}	{1}
Car	-	-	-	-	{0}	{1}	-	-	-	-	-
D_A	-	-	{1}	{1}	{2,9}	{2,9}	{2,9}	{3,10}	{3,10}	{4}	{4}
#Hol	-	-	-	-	(<-,3]	(<-,3]	(3,->)	-	-	-	-
Age	-	-	-	-	-	-	-	(<-,48]	(48,->)	-	-
Hotel	.4324	.0938	.8448	.4286	.8158	.9344	.7347	.2712	.6452	.6129	.8354
Nperm	.0811	.0104	.0345	.0204	.0000	.0000	.0000	.0339	.0000	.0161	.0253
Perm	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Other	.4865	.8958	.1207	.5510	.1842	.0656	.2653	.6949	.3548	.3710	.1393
N	37	96	58	49	38	122	49	59	62	62	79
Rule#	11	12	13	14	15	16	17	18	19	20	21

Table 11.1 Continued - Accommodation

#Tper	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Fracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Chld	{1}	{1}	{1}	{1}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}
Lngth	-	-	-	-	(<-,2]	(2,->)	-	-	-	-
D_A	{5,8}	{6}	{7}	{11}	{1,7,8,10}	{1,7,8,10}	{2,9}	{3}	{3}	{4}
Cst	-	-	-	-	-	-	-	{1}	{2,3,4,5,6}	-
Hotel	.8587	.9057	.6970	.7407	.6849	.1294	.8723	.3243	.0784	.3043
Nperm	.0054	.0000	.0455	.0074	.0137	.0235	.0000	.0000	.0000	.0580
Perm	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Other	.1359	.0943	.2575	.2519	.3014	.8471	.1277	.6757	.9216	.6377
N	184	106	66	135	73	85	94	37	51	69
Rule#	22	23	24	25	26	27	28	29	30	31

When *no holidays* have been spent *in the personal accommodation with a permanent location*, the *possession* of such an accommodation itself is important. *People who possess a tourist accommodation with a permanent location* will often select this alternative for *domestic holidays* (R74-78; except for short breaks “alone” or with “adults only” (R74)). For *foreign holidays*, however, the possession of a permanent personal accommodation does not automatically lead to the selection of this alternative (possibly because the accommodation is in the Netherlands).

Table 11.1 Continued - Accommodation

#Tper	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Fracc	{0}	{0}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
D_NLA	{2}	{2}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Chld	{2,3}	{2,3}	-	-	-	-	-	-	-	-
Lngth	-	-	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]
Party	-	-	{1,3,4,99}	{2}	{2}	{5}	{1,2}	{1,2}	{3}	{4,5}
D_A	{5}	{6,11}	-	-	-	-	-	-	-	-
#Dnper	-	-	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)	(0,->)	(0,->)
D_NL	-	-	-	{1,3,4,8}	{2,5,6,7,9}	-	-	-	-	-
#DOA	-	-	-	-	-	-	(<-,0]	(0,->)	-	-
Seas+	-	-	-	-	-	-	-	-	-	{1,6}
Hotel	.7447	.6250	.1827	.8226	.5158	.0886	.4154	.2000	.0317	.2162
Nperm	.0213	.0000	.1154	.0000	.1667	.0633	.4000	.2200	.0541	.0541
Perm	.0000	.0000	.0096	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Other	.2340	.3750	.6923	.1774	.3175	.8481	.1846	.5800	.9142	.7297
N	47	56	104	62	54	79	65	50	37	37
Rule#	32	33	34	35	36	37	38	39	40	41

Table 11.1 Continued - Accommodation

#Tper	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Fracc	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Lngth	(<-,1]	(<-,1]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]
Party	{4,5}	{99}	{1,2,99}	{3,4,5}	{3,4,5}	-	{1,2,3,99}	{1,2,3,99}	{4,5}
#Dnper	(0,->)	(0,->)	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)	(0,->)	(0,->)
Seas+	{2,3,4,5}	-	-	-	-	{1,6}	{2,3,4,5}	{2,3,4,5}	{2,3,4,5}
Sclss	-	-	-	(<-,2]	(2,->)	-	-	-	-
Prov	-	-	-	-	-	-	{1,2,4,9,12}	{3,5,6,7,8,10,11}	-
Hotel	.0462	.0652	.2273	.0962	.0000	.0488	.0000	.0577	.0000
Nperm	.3231	.8478	.1818	.0962	.2281	.1707	.5000	.8846	.4375
Perm	.0000	.0000	.0303	.0000	.0000	.0000	.0278	.0000	.0000
Other	.6307	.0870	.5606	.8076	.7719	.7805	.4722	.0577	.5625
N	65	46	66	104	57	41	36	52	48
Rule#	42	43	44	45	46	47	48	49	50

Under these circumstances, hotels are very popular for *holidays up till 15 days* to Spain and Portugal, Austria and Switzerland, Italy and Greece, and South East Mediterranean countries (R80). For holidays of this duration to other countries the tourist accommodations that are not owned by the tourist are preferred (R79). *Longer foreign holidays*, finally, are more frequently spent in (permanent or non-permanent) personal accommodations (R81).

Table 11.1 Continued - Accommodation

#Tper	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Fracc	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
D_NLA	{1}	{1}	{1}	{1}	{2}	{2}	{2}	{2}	{2}	{2}
Lngth	(2,->)	(2,->)	(2,->)	(2,->)	(<-,2]	(2,3]	(2,3]	(2,3]	(3,->)	(3,->)
D_A	-	-	-	-	-	{1,3,4,7}	{2,5,9,11}	{6,8,10}	-	-
#Hol	(<-,1]	(1,3]	(1,3]	(3,->)	-	-	-	-	-	-
Seas+	-	-	-	-	-	-	-	-	{1,5,6}	{2}
Urban	-	(<-,3]	(3,->)	-	-	-	-	-	-	-
#Thot	-	-	-	-	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Prio4	-	-	-	-	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]
Hotel	.0556	.0000	.0000	.0000	.5526	.0909	.8077	.5366	.5652	.1781
Nperm	.6944	.5122	.7586	.7568	.1316	.5051	.0256	.1220	.1304	.5890
Perm	.0139	.0000	.0811	.0811	.0000	.0000	.0000	.0000	.0000	.0000
Other	.2361	.4878	.1603	.1621	.3158	.4040	.1667	.3414	.3044	.2329
N	72	82	58	74	76	99	78	41	46	73
Rule#	51	52	53	54	55	56	57	58	59	60

When *people do not own a permanent personal accommodation*, the possession of a non-permanent personal tourist accommodation and the destination of the holiday are the most important conditions. *Foreign holidays of people who own a non-permanent tourist accommodation* are often spent in hotels if people have already done this during holidays with higher profiling priorities (65.38%; R70-73). If people have not planned any hotel-based holidays yet, hotels are still the most preferred alternative (35.59%), but under these conditions the personal non-permanent accommodation gains importance (34.59%; R55-69), especially when the holiday with the highest profiling priority is a long or extra long holiday (R60-63), or a medium holiday to one of the neighbouring countries (R56). “Other” accommodation types that are not owned by the tourist are the most preferred option for the second and third most important holidays within the annual trip program to France and Germany (when no holidays have been planned for this accommodation type yet; R64) and to Belgium and Luxembourg and Scandinavia and Denmark (R67).

Table 11.1 Continued - Accommodation

#Tper	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Fracc	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Lngth	(3,->)	(3,->)	(3,->)	-	-	-	-	-	-	-
D_A	{1,3,5, 7,8,10}	{1,3,5,7, 8,10}	{2,4,6, 9,11}	{1,4}	{1,4}	{2,5, 8,9}	{3,10}	{6,7, 11}	-	{1,4}
#Dnper	-	-	-	(<-,0]	(0,->)	-	-	-	-	-
Seas+	{3,4}	{3,4}	{3,4}	-	-	-	-	-	-	-
Sclss	(<-,1]	(1,->)	-	-	-	-	-	-	-	-
#Thot	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)
Prio4	(<-,1]	(<-,1]	(<-,1]	(1,3]	(1,3]	(1,3]	(1,3]	(1,3]	(3,->)	-
Gndr	-	-	-	-	-	-	-	-	-	{1}
Hotel	.0882	.0061	.2020	.4600	.4478	.7553	.1563	.6441	.4571	.7556
Nperm	.6324	.7561	.4343	.0200	.2836	.1277	.2500	.0169	.2857	.1111
Perm	.0000	.0000	.0000	.0200	.0000	.0000	.0000	.0000	.0000	.0000
Other	.2794	.2378	.3637	.5000	.2686	.1170	.5937	.3390	.2572	.1333
N	68	164	99	50	67	94	64	59	35	45
Rule#	61	62	63	64	65	66	67	68	69	70

Domestic holidays of people who own a non-permanent tourist accommodation, are spent in that accommodation most often when the holiday under consideration lasts *at least 9 days* (R51-54). For *medium long holidays*, the “other” accommodation types are often preferred when no holidays have been planned for the personal non-permanent accommodation yet (R44-47). When, however, at least one holiday has been planned for the personal non-permanent accommodation, this alternative is the most favourite alternative for subsequent holidays (R48-50), especially during the spring and summer season (R49+50). For *domestic short breaks by people who own a non-permanent tourist accommodation*, finally, accommodations that are not owned by the tourist are preferred when no days have been planned for the own non-permanent accommodation yet (R34-37). When at least one day has been spent in the own non-permanent accommodation, the “other” accommodation types that are not owned by the tourist are selected more frequently (R38-43). Exceptions occur, however, when no days have been planned previously for this accommodation type, in which case both hotels and the own non-permanent accommodation are equally preferred (R38). Also, the personal non-permanent accommodation is preferred when the travel party is unknown (indicating that this holiday is the third or higher holiday within a quarter; R43).

Table 11.1 Continued - Accommodation

#Tper	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]
Peracc	{0}	{0}	{0}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Fracc	{1}	{1}	{1}	-	-	-	-	-	-	-
D_NLA	{2}	{2}	{2}	{1}	{1}	{1}	{1}	{1}	{2}	{2}
Lngth	-	-	-	(<-,2]	(<-,2]	(2,3]	(3,4]	(4,->)	(<-,3]	(<-,3]
Party	-	-	-	{1,2}	{3,4,5,99}	-	-	-	-	-
D_A	{1,4}	{2,5,6,7, 8,9,11}	{3,10}	-	-	-	-	-	{1,3,4,7, 8,10,11}	{2,5, 6,9}
#Thot	(0,->)	(0,->)	(0,->)	-	-	-	-	-	-	-
Gndr	{2}	-	-	-	-	-	-	-	-	-
Hotel	.4773	.8037	.4531	.3333	.0938	.0233	.0000	.0000	.3529	.8636
Nperm	.1818	.0093	.0781	.1667	.0313	.1628	.0417	.0000	.1324	.0000
Perm	.0000	.0000	.0000	.0278	.4688	.7209	.9583	1.000	.1324	.0455
Other	.3409	.1870	.4688	.4722	.4061	.0930	.0000	.0000	.3823	.0909
N	44	107	64	36	64	43	48	38	68	44
Rule#	71	72	73	74	75	76	77	78	79	80

For *foreign holidays by people who do not own any personal accommodations (neither permanent nor non-permanent)*, the presence of children in the household is important in the choice of accommodation (R13-33). *Households without children* usually go to hotels (R13-25). Other accommodation types that are not owned by the tourist are preferred during short breaks to France (R14) and holidays to Belgium, Luxembourg, Scandinavia and Denmark by people younger than 49 years (R18). *Households with children* prefer the available accommodation categories that are not owned by themselves for their foreign holidays (26-33). However, hotels are more desirable during holidays to Spain and Portugal, Austria and Switzerland, Italy and Greece, South East Mediterranean countries and the other parts of the world outside Europe (R28+32+33), and for short breaks to France, the UK and Ireland, the (former) Eastern European countries and Scandinavia and Denmark (R26). In contrast, for holidays by these households to Belgium, Luxembourg and Germany and longer holidays to France, the UK and Ireland, (former) Eastern European countries and Scandinavia and Denmark, the “other” accommodation types are preferred (R27+29-31).

Table 11.1 Continued - Accommodation

#Tper	(<-,0]	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)	(0,->)
Peracc	{1}	-	-	-	{0}	{1}	{1}	-	-	-
D_NLA	{2}	{1}	{1}	{1}	{1}	{1}	{1}	{2}	{2}	{2}
Lngth	(3,->)	-	-	-	-	-	-	-	-	-
Party	-	{1,4}	{2,3,5}	{2,3,5}	{99}	{99}	{99}	{1,99}	{2,3,4,5}	{2,3,4,5}
HolP	-	-	-	-	-	-	-	-	{0}	{1}
Seas+	-	-	{1,3,6}	{2,4,5}	-	-	-	-	-	-
Ski	-	-	-	-	-	{0}	{1}	-	-	-
Hotel	.1940	.0000	.3000	.0926	.0227	.0000	.0000	.0213	.8000	.5870
Nperm	.3134	.0500	.0250	.1481	.0909	.0000	.0000	.0000	.0000	.1522
Perm	.2687	.9000	.1250	.5000	.8864	1.000	.9737	.9362	.1500	.1522
Other	.2239	.0500	.5500	.2593	.0000	.0000	.0263	.0425	.0500	.1086
N	67	40	40	54	44	775	38	47	40	46
Rule#	81	82	83	84	85	86	87	88	89	90

For *domestic holidays by people who do not own a personal tourist accommodation*, finally, the presence of children in the household is important too (R1-12). The majority of short breaks by tourists from *households without children* are spent in hotels (R 1-4), except when the travel party includes children and/or comprises at least 9 people (R4). Domestic holidays of a minimum 5 days by these households are preferably spent in “other” accommodations (R5+6). *Domestic holidays of households with (schoolgoing or other) children*, finally, are also often spent in “other” accommodations (R7-12).

11.2.5 Statistics and validation

Table 11.2 presents the confusion matrix and statistics of the DT for the choice of accommodation that was discussed in the previous section. The DT is able to correctly classify 64.60% of all observed holidays, a tremendous improvement compared to the 28.91% of the probabilistic null-model. Holidays to accommodations with permanent locations that are owned by the tourist are correctly predicted in 90.95% of the cases, which should be attributed to the fact that people owning such an accommodation often select this alternative. Although not as spectacular, for the other accommodation alternatives too, the improvement of the predictive power is significant. For “hotel” and “other” the predictive accuracy almost doubles compared to the null-model, whereas for “non permanent accommodations owned by the traveller” the number of hits increases by 243.4% to 48.11% of the observations. The excellent performance of the DT is also evident from the contingency coefficient that amounts to 0.7878, where the maximum for a 4×90 DT is 0.8660.

Table 11.2 Confusion matrix & statistics for the DT *accommodation*

	Hotel’	Nper’	Per’	Other’	Total’
Hotel	**.6364**	.06132	.0122	.2901	.3296
Nper	.1449	**.4811**	.0353	.3387	.1401
Per	.0259	.0280	**.9065**	.0396	.1643
Other	.2603	.1219	.0174	**.6004**	.3659
Total	.3296	.1401	.1643	.3659	**.6460**

N° of observations	7121	
Stopping criteria	80 before/35 after	
α	0.05	
N° of columns (= rules)	90	
Theta (1 col)	2606 (.3659)[a]	2059 (.2891)[b]
Theta (90 cols)	5278 (.7412)[a]	4600 (.6460)[b]
χ^2_{DT}(df) / Cont. coef.	11609.18 (267) / .7878	

[a] Deterministic [b] Probabilistic

11.3 The question of which mode to travel by

11.3.1 The choice sets

Travelling, by definition, is an essential part of tourism. Tourism mobility has increased tremendously over the past decades, and the car has become the dominant means of transport for leisure purposes (Stemerding, 1996). In addition, the aeroplane has become very popular for destinations abroad. Figure 11.1 displays the present model split for holidays by the Dutch population in the Netherlands and abroad.

Simultaneously, however, travelling has also become one of the most problematic aspects of tourism because cars are an important source of congestion and crowding. In addition, cars and aeroplanes seriously contribute to environmental and noise pollution. Research among Dutch tourists who had made a foreign holiday in Europe during the summer of 1996 or 1997, showed that environmental considerations were rarely taken into account in the choice of transport mode. Rather, the deciding factors in the mode choices included price, comfort, party size (for families), privacy, safety, accessibility and travel time (for long distance holidays). Also, it was concluded that one of the main problems for train- and coach-tours is the negative image of these modes. Many people think these options to be rather slow, uncomfortable and expensive (Wasser, 1998).

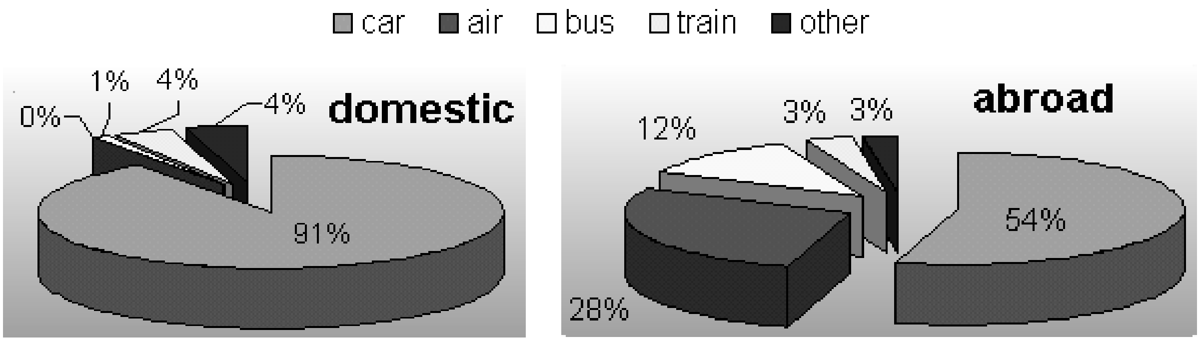


Figure 11.1 Model split for domestic and foreign holidays by the Dutch population in 1998 (source: 1998 CVO-data)

In order to identify the factors that condition the choice between transport modes that are harmful to the environment and those that are (more) ecologically sound, the *MERLIN*-system first considers the choice between “car”, “aeroplane” or an “alternative” means of transport. In this choice set, all alternative modes, including buses and coaches, trains (including “car-sleep-trains”), boats and ferries, walking and cycling, and other means public transport (such as trams and subways), are grouped together for both conceptual and methodological reasons. Conceptually, these modes are often considered as (more) ecologically sound, thus distinguishing themselves from the cars and aeroplanes. Methodologically, the CHAID-based algorithm may not be able to identify decision rules that identify these alternative modes separately because of the unbalanced response distribution over the alternatives (in 1998, 74.9% of all holidays were made by car, 12.2% by aeroplane and only 12.9% by all other modes together). Given the policy and industry interests in advancing bus- and train-based holidays, the *MERLIN*-system will further examine the factors that condition the choice of the “alternative” mode. More specifically, given the choice to use an alternative mode, the factors that condition the preferences for the “bus”, the “train” and all “other” alternatives will be examined. Transport mode choices for holidays will thus be described using a hierarchy of two DTs.

11.3.2 The condition variables

In addition to the condition variables common to all choice facets (Table A1.1 in Appendix 1) and the choices that have been considered previously, i.e the travel party, the three conditions related to the profiling priority of the holiday, the decision whether or not to travel during a school holiday period, and the choices of season, destination and accommodation (see Table A1.2), several condition variables are common to both transport mode DTs (see Table A1.17). First, as before, the MERCIN-system assumes the total amount of expenditures that have already been planned for holidays with higher profiling priorities to condition mode choices, because financial budgets may induce the tourist to select cheaper modes when (several) expensive holidays have already been planned.

Second, the variables that represent the tourist's propensity to select the respective transport mode alternatives, including car, aeroplane and "alternative" modes, are included. Recall that these propensities were calculated using the MNL-model discussed in section 10.2, and included in the DTs for the destination choices to make these choice rules susceptible to general changes in transport mode preferences.

Third and finally, several condition variables describe the choices regarding transport modes for previously considered holidays both in the terms of the number of trips and the number of days that are involved with these choices. Six of these variables summarise the number of car- and air-based holidays and the number of holidays that are based on alternative modes (and the days that are associated with these holidays). These six variables are common to all transport mode DTs. Finally, there are six condition variables that are specific to the choice between the alternative modes. These variables summarise the pervious choices for the three alternative mode options, i.e. bus, train and "other" both in terms of the numbers of trips and the number of days that are involved (Table A1.18).

11.3.3 The stopping criteria of the CHAID-based algorithm

The final decision regarding the DTs for tourist transport mode choices comprises the selection of the stopping criteria of the CHAID-based algorithm. For each transport DT, the α-level for predictor eligibility is set at 5%. With regard to the minimum number of observations before/after splitting the sample, the performance of various models is reported in Appendix 2 (Tables A2.15 and A2.16). Based on these analyses, the stopping criteria for the choice between car-, air- and alternative modes are set at 60 observations before and 25 observations after the division of a decision rule because the resulting set of decision rules performs best on both the cross-validation and the aggregate measures. For similar reasons, the DT for the choice of the alternative transport modes is induced using a minimum of 35 observations before and 15 observations after division.

11.3.4 The generated decision tables

The first transport DT ("main mode of transport") consists of 100 decision rules that describe the factors that condition the choices between the modes "car", "aeroplane" ("Air" in the DT) and "alternative" ("Alter" in the DT). The DT is presented in Table 11.3.

Table 11.3 DT for the choice of the main mode of transport

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Car	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
#Tcar	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)	(0,->)
#Dalt	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,->)	(0,->)	-	-	-
Csize	(<-,4]	(<-,4]	(4,5]	(5,->)	(5,->)	-	-	-	-	-
Cst	{1,2,3,5}	{4,6}	-	-	-	-	-	-	-	-
Region	-	-	-	{1,2}	{3}	-	-	-	-	-
Acc	-	-	-	-	-	(<-,3]	(3,->)	-	-	-
WrkW	-	-	-	-	-	-	-	(<-,0]+{9}	(<-,0]+{9}	(0,->)\{9}
Party	-	-	-	-	-	-	-	{1,3,4,99}	{2,5}	-
Car	.6410	.9318	.4211	.5000	.8387	.0714	.3571	.9756	.7838	.6567
Air	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Alter	.3590	.0682	.5789	.5000	.1613	.9286	.6429	.0244	.2162	.3433
N	39	44	38	54	31	28	42	82	37	67
Rule#	1	2	3	4	5	6	7	8	9	10

Evidently, the most important condition is the destination in terms of domestic and abroad. This is probably due to the fact that *domestic destinations* are never visited by air (R1-41). The majority of domestic holidays are made by car (90.7%). Even for 66.02% of all domestic holidays by tourists from *households that do not own a car*, this mode is selected (R1-10). Apparently, these tourists borrow or rent cars or they travel together with other people. The probability of selecting alternative modes only exceeds 50% when the tourist has already made holidays with alternative travel modes (R6+7). This probability equals or exceeds 50%, when the tourist has not made any other holidays by car or alternative mode, and the tourist lives in a city with 50 to 100 thousand residents (R3), or in a city in the mid or southern holiday region with 100 to 250 thousant residents (R4).

Table 11.3 Continued - main mode of transport

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Car	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
#Dalt	(<-,0]	(<-,0]	(<-,0]	(0,->)	-	-	-	-	-	-
Region	-	-	-	-	-	-	{1,3}	{2}	-	-
Party	{1,2}	{1,2}	{1,2}	{1,2}	{3}	{3}	{4}	{4}	{5}	{5}
D_NL	{1,4,7}	{1,4,7}	{1,4,7}	{1,4,7}	{1,4,7}	{1,4,7}	{1,4,7}	{1,4,7}	{1,4,7}	{1,4,7}
Gndr	{1}	{2}	{2}	-	-	-	-	-	-	-
Lngth	-	(<-,2]	(2,->)	-	(<-,3]	(3,->)	-	-	-	-
Work	-	-	-	-	-	-	-	-	{0,3}	{1,2}
Car	.9343	1.000	.9375	.8837	.9952	.9286	1.000	.9048	.7125	.9306
Air	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Alter	.0657	.0000	.0625	.1163	.0048	.0714	.0000	.0952	.2875	.0694
N	137	128	32	43	209	28	108	42	80	72
Rule#	11	12	13	14	15	16	17	18	19	20

As expected, *tourists from households with at least one car* chose to travel by car even more frequently for domestic holidays than their counterparts without a car (on average, 94.11% across R11-41). The probability of selecting an alternative transport mode is relatively high when the travel party consists of at least nine people, the destination is Utrecht, ‘t Gooi and the Veluwe (Land North-mid), the North Sea coastal area, or the Western and Mid parts of North-Brabant (Land South-west), and the tourist is younger than 16 years old and/or has no paid job (R19).

Table 11.3 Continued - main mode of transport

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Car	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
#Tcar	-	-	-	(<-,0]	(<-,0]	(<-,0]	(<-,0]	(0,3]	(0,3]	(3,->)
#Dalt	(<-,0]	(<-,0]	(0,->)	-	-	-	-	-	-	-
Party	{99}	{99}	{99}	-	-	-	-	-	-	-
D_NL	{1,4,7}	{1,4,7}	{1,4,7}	{2,9}	{2,9}	{2,9}	{2,9}	{2,9}	{2,9}	{2,9}
Peracc	{0}	{1}	-	-	-	-	-	-	-	-
Prov	-	-	-	{1,2,5, 9,12}	{1,2,5, 9,12}	{3,4,6,7, 8,10,11}	{3,4,6,7, 8,10,11}	-	-	-
Fracc	-	-	-	{0}	{1}	-	-	-	-	-
Inc	-	-	-	-	-	(<-,6]	(6,->)	-	-	-
Cst								{1,2,3,4,5}	{6}	
Car	.9722	1.000	.8372	.5660	.7879	.9524	.7674	.9259	.7742	.9897
Air	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Alter	.0278	.0000	.1628	.4340	.2121	.0476	.2326	.0741	.2258	.0103
N	36	447	43	53	33	84	43	189	62	97
Rule#	21	22	23	24	25	26	27	28	29	30

Table 11.3 Continued - main mode of transport

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Car	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Acc	-	-	-	-	-	-	-	-	{1,2}	{3,4}
Party	{1,5}	{2,4}	{2,4}	{3,99}	{3,99}	-	-	-	-	-
D_NL	{3,8}	{3,8}	{3,8}	{3,8}	{3,8}	{5,6}	{5,6}	{5,6}	{5,6}	{5,6}
Work	-	-	-	{0,1,3}	{2}	-	-	-	-	{1,2,3}
Sclss	-	(<-,3]	(3,->)	-	-	-	-	-	-	-
Age	-	-	-	-	-	(<-,8]	(8,17]	(17,29]	(29,->)	(29,->)
Car	.7529	.9005	1.000	.9945	.9091	.9921	.8947	.8286	.9709	1.000
Air	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Alter	.2471	.0995	.0000	.0055	.0909	.0079	.1053	.1714	.0291	.0000
N	85	221	79	181	33	127	76	35	172	384
Rule#	31	32	33	34	35	36	37	38	39	40

Second, this probability is also higher when the holiday destination is the northern water areas (Wadden Islands, the beaches of the Lake Yssel and the lake areas in Groningen, Friesland and North-West Overijssel) or the “other” tourist areas (the cities and the remaining parts of the Netherlands) (R24-30). This is probably explained best by the fact that cars are prohibited on some of the Wadden Isles. Also, the costs of transporting one’s car to the isles may be relevant. With regard to the four cities, finally, the higher popularity of alternative modes should probably be attributed to the fact that these urban areas are reasonably accessible by public transport.

Table 11.3 Continued - main mode of transport

D_NLA	{1}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Car	{1}	-	-	-	-	-	-	-	-	-
Region	-	-	-	-	-	-	-	-	-	{1,2}
Acc	{3,4}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
D_NL	{5,6}	-	-	-	-	-	-	-	-	-
Lngth	-	(<-,1]	(<-,1]	(<-,1]	(1,2]	(1,2]	(1,2]	(1,2]	(2,3]	(2,3]
Work	{0}	-	-	-	-	-	-	-	-	-
Fracc	-	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Age	(29,->)	-	-	-	-	-	-	-	-	-
Chld	-	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
D_A	-	{1}	{2,5,6,7, 9,10}	{3,4,8,11}	{1,7,8,10}	{2,6,9,11}	{3,5}	{4}	{1,7,8,10}	{2}
Car	.9706	.3793	.0345	.6786	.2931	.0000	.5965	.4808	.2203	.0000
Air	.0000	.0690	.5862	.0119	.1552	1.000	.0526	.0192	.1356	.7838
Alter	.0294	.5517	.3793	.3095	.5517	.0000	.3509	.5000	.6441	.2162
N	34	29	29	84	58	76	57	52	59	37
Rule#	41	42	43	44	45	46	47	48	49	50

Table 11.3 Continued - main mode of transport

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Region	{3}	-	-	-	-	-	-	-	-	-
Acc	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{2,3}	{4}
Lngth	(2,3]	(2,3]	(2,3]	(2,3]	(2,3]	(3,->)	(3,->)	(3,->)	-	-
Fracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
D_A	{2}	{3,4,5}	{6}	{9}	{11}	{1,4,5,7,8,10}	{2,6}	{3,9,11}	-	{1,5}
Car	.0385	.6316	.0615	.0000	.0000	.8000	.2167	.0000	.8444	.9107
Air	.3846	.0132	.7846	.9643	1.000	.0000	.5167	.9867	.0444	.0357
Alter	.5769	.3552	.1539	.0357	.0000	.2000	.2666	.0133	.1112	.0536
N	26	76	65	28	31	35	60	75	45	56
Rule#	51	52	53	54	55	56	57	58	59	60

However, the probability of leaving the car at home only exceeds 20% when the domestic destination is one of the water areas in the North or the “other parts of the Netherlands”, *and* when a person with the civil state “other” (usually children) has made only 1 to 3 car-based holidays (R29), or, when a person has not made any car-based holidays yet, lives in Groningen, Friesland, Flevoland, South-Holland or Limburg (R24+25) or in one of the other provinces and has an annual household income of at least € 25,000 (R27). Finally, the probability of selecting an alternative means of transport amounts to 24.71% when the tourist visits water areas in the mid of the Netherlands or the (south-) eastern parts of North-Brabant and Gelderland or the province of Limburg (Land South-south) either alone or with a party of at least 9 people (R31).

Table 11.3 Continued - main mode of transport

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Acc	{4}	{4}	{4}	{4}	{4}	{1}	{1}	{1}	{1}	{1}
Work	-	-	-	-	-	-	{0,1}	{2,3}	-	-
Fracc	{0}	{0}	{0}	{0}	{0}	{1}	{1}	{1}	{1}	{1}
Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
D_A	{2,6,9}	{3}	{4}	{7,8,10}	{11}	{1}	{2,6,9}	{2,6,9}	{3}	{4}
Car	.3415	.8333	.7500	.4565	.0000	.6250	.0270	.0784	.8788	.8529
Air	.6098	.0417	.0000	.2826	1.000	.0313	.9189	.6667	.0000	.0000
Alter	.0487	.1250	.2500	.2609	.0000	.3437	.0541	.2549	.1212	.1471
N	41	48	36	46	37	64	74	51	33	34
Rule#	61	62	63	64	65	66	67	68	69	70

Evidently, for *foreign holidays*, aeroplanes and alternative modes of travel gain importance, but cars still dominate the travel mode choices because 56.25% of all foreign destinations are reached by car (R42-100). In the travel mode choices for foreign holidays, the presence of children in the household is the most significant condition. For *household without children*, the next most important condition is the possession of a *tourist accommodation without a permanent location* (R42-81), where tourists who own such an accommodation are more inclined to travel by car (R66-81). However, the aeroplane is preferred when people plan a hotel-based holiday to Spain and Portugal, Italy and Greece, South East Mediterranean countries or “other” parts of the world (R67+68+73), or a holiday in an “other” accommodation to these destinations (except for Italy and Greece; R78). Aeroplane and alternative modes of travel are equally preferred when people plan a hotel-based holiday to the UK, Ireland, (former) Eastern European countries, Scandinavia or Denmark (R72).

Table 11.3 Continued - main mode of transport

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Acc	{1}	{1}	{1}	{2,3}	{2,3}	{2,3}	{4}	{4}	{4}	{4}
Fracc	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
D_A	{5}	{7,8,10}	{11}	{1,4,8,9,10}	{2,6,11}	{3,5,7}	{1}	{2,9,11}	{3}	{4,5}
Car	.8043	.2157	.0000	.9784	.8621	.8600	.9333	.2041	.8929	.9615
Air	.0000	.3922	1.000	.0072	.1379	.0000	.0000	.7143	.0000	.0000
Alter	.1957	.3921	.0000	.0144	.0000	.1400	.0667	.0816	.1071	.0385
N	46	51	51	139	29	50	45	49	28	26
Rule#	71	72	73	74	75	76	77	78	79	80

When *households without children do not own a tourist accommodation without a permanent location*, the choice of accommodation also plays an important role in the choice of transport mode (R42-65). When these households do not plan to stay in a hotel, car is by far the most preferred option (R59-65). This does not apply, however, when the holiday is planned to Spain and Portugal, Italy, Greece, South East Mediterranean countries or other parts of the world outside Europe. For for these destinations, the aeroplane is often preferred (R61+65). When these households plan a *hotel-based holiday*, the duration and the destination of the holiday are important (R42-58). For *long and extra long holidays*, the aeroplane is often chosen (R57-58), except when the destination is France, Germany, Austria and Switzerland, the UK and Ireland, (former) Eastern European countries or Scandinavia and Denmark (R56). For *extended holidays* too, the aeroplane is often preferred (R49-55). However, alternative modes are selected more frequently when the destination is France, the UK, Ireland, (former) Eastern European countries, Scandinavia, Denmark, Spain or Portugal (R49+51), whereas the car is preferred for holidays to Belgium, Luxembourg, Germany, Austria and Switzerland (R52).

Similarly, *medium long hotel-based holidays* to these five countries, too, are often by car or alternative modes of transport (R47-48), whereas for holidays of this duration to Spain and Portugal, Italy and Greece, South East Mediterranean countries and other parts of the world outside Europe, the aeroplane is always the preferred means of transport (R46). Medium long holidays to France, UK and Ireland, (former) Eastern European countries or Scandinavia and Denmark are often planned using alternative means of transport (R45). For *hotel-based short breaks* by these households, finally, holidays to France are often planned by alternative travel mode or car (R42; this may be due to the popular 3- or 4-day organised bus- or train-trips to Paris and/or Eurodisney). Short breaks to Belgium, Luxembourg, (former) East European countries, and other parts of the word[35], however, are often by car (R44), while short breaks to all other destinations are often by aeroplane (R43).

Table 11.3 Continued - main mode of transport

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Acc	{4}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Lngth	-	-	(<-,2]	(2,->)	(2,->)	-	-	-	-	-
Fracc	{1}	-	-	-	-	-	-	-	-	-
Sclss	-	-	-	(<-,3]	(3,->)	-	-	-	-	-
Chld	{1}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}
D_A	{6,7,8,10}	{1}	{2}	{2}	{2}	{3,5}	{4}	{6}	{7,8,10}	{9}
Car	.5192	.5632	.0385	.0769	.1923	.8269	.9091	.0811	.1957	.0357
Air	.1731	.0230	.9615	.6346	.2308	.0096	.0000	.7838	.2609	.8929
Alter	.3077	.4138	.0000	.2885	.5769	.1635	.0909	.1351	.5434	.0714
N	52	87	26	52	26	104	44	37	46	28
Rule#	81	82	83	84	85	86	87	88	89	90

For *households with children (both schoolgoing and other) planning a foreign holiday* too, the planned accommodation type is the most significant condition. Again, when these households plan to use their permanent or non-permanent accommodation or one of the “other” accommodations, car is by far the most preferred option (R92-100). Under these conditions, exceptions occur when the accommodation is “other” and the destination is France or Scandinavia and/or Denmark, because under these conditions, the aeroplane is preferred (R95). Similarly, when, the intended destination is Italy and Greece, the UK and Ireland, or one of the (former) Eastern European countries, cars and alternative travel

[35] This is probably explained by the lack (or low number) of observed short breaks to “other parts of the world”. This may produce seemingly strange decision.

modes are almost equally preferred (R100). Finally, when *households with children plan a hotel-based foreign holiday*, the destination is important (R82-91). Under these conditions, cars and alternative means of transport are favourite for holidays to France (R82), whereas the car is preferred for Belgium, Luxembourg, Germany, Austria and Switzerland (R86+87). Spain and Portugal are often reached by air (R83-85; except for holidays of at least 9 days by people from the lowest two social classes (R85)), which also applies to Italy and Greece, South East Mediterranean countries and other parts of the world outside Europe (R88+90+91). Alternative travel modes are preferred for hotel-based holidays by households with children to the UK and Ireland, (former) East European countries and Scandinavia and Denmark (R89).

Table 11.3 Continued - main mode of transport

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Acc	{1}	{2,3}	{4}	{4}	{4}	{4}	{4}	{4}	{4}	{4}
Gndr	-	-	{1}	{2}	-	-	-	-	-	-
Fracc	-	-	-	-	-	{0}	{1}	-	-	-
Chld	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}
D_A	{11}	-	{1,10}	{1,10}	{2,9,11}	{3}	{3}	{4,5}	{6,7,8}	{6,7,8}
#Hol	-	-	-	-	-	-	-	-	(<-,2]	(2,->)
Car	.0000	.9790	.9176	.9865	.3902	.9867	.9153	.9596	.7949	.4000
Air	1.000	.0070	.0000	.0000	.5610	.0000	.0000	.0000	.1282	.1600
Alter	.0000	.0140	.0824	.0135	.0488	.0133	.0847	.0404	.0769	.4400
N	25	286	85	74	41	75	59	99	39	25
Rule#	91	92	93	94	95	96	97	98	99	100

Based on 918 observations, 32 decision rules for the choices between the alternative transport modes were induced. The DT is presented in Table 11.4. Again the difference between transport mode choices for domestic and foreign holidays is eminent. *Domestic destinations* are rarely reached by bus (R1-14). Only when the travel party consists of *at least 9 people*, this mode is selected in more than 20% of the cases (R12-14). When the *travel party is unknown or includes children*, "other" modes of transport are often preferred (R7-11), except when people do not own a car and live in a highly urbanised area (R7), or when they live in other areas and plan to visit the metaregions Land North (except for the Land North–Mid), Land South or "Other", or the North Sea coastal area (R9). Finally, for domestic holidays with alternative means of transport by people travelling by themselves or with adults only, the train is often preferred, but the difference with the "other" alternative is small (55.32% vs. 40.43%; R1-6).

Table 11.4 DT for the choice of the alternative transport mode

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Party	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{1,2}	{3,4,99}	{3,4,99}	{3,4,99}
Age	(<-,42]	(<-,42]	(42,55]	(42,55]	(55,64]	(64,->)	-	-	-
Region	{1}	{2,3}	-	-	-	-	-	-	-
Car	-	-	{0}	{1}	-	-	{0}	{0}	{0}
Urban	-	-	-	-	-	-	(<-,1]	(1,->)	(1,->)
D_NL	-	-	-	-	-	-	-	{1,2,3}	{4,5,6,7,8,9}
Bus	.0455	.0000	.0000	.0000	.0000	.1786	.0000	.0000	.0000
Train	.3182	.7742	.3182	.1176	.7826	.7143	.9474	.0667	.8333
Other	.6363	.2258	.6818	.8824	.2174	.1071	.0526	.9333	.1667
N	22	31	19	17	23	28	19	15	24
Rule#	1	2	3	4	5	6	7	8	9

For *foreign holidays*, the bus is by far the most popular alternative means of transport (67.91%; R15-32). Under these conditions, the accommodation is key. For foreign holidays to or with the *personal accommodation (with or without a permanent location)*, "other" modes are often used (R28). In the alternative mode choice for foreign holidays with *"other" accommodations*, the destination is important, where UK and Ireland are often reached by "other" modes (R32), while other destinations are visited more frequently by bus (R29-31). However, when the holiday under consideration has the highest profiling priority within the annual tourist trip program and the destination is Belgium, Luxembourg, Germany, Scandinavia, Denmark or one of the other parts of the world outside Europe, both the train and the "other" alternatives modes are almost equally preferred (R30).

Table 11.4 Continued - alternative transport mode

D_NLA	{1}	{1}	{1}	{1}	{1}	{2}	{2}	{2}
Party	{3,4,99}	{3,4,99}	{5}	{5}	{5}	-	-	-
Car	{1}	{1}	-	-	-	-	-	-
Urban	-	-	(<-,2]	(<-,2]	(2,->)	-	-	-
D_NL	-	-	{1,2,3,8}	{4,5,6,7,9}	-	-	-	-
Peracc	{0}	{1}	-	-	-	-	-	-
Acc	-	-	-	-	-	{1}	{1}	{1}
D_A	-	-	-	-	-	{1}	{1}	{1}
#Tcar	-	-	-	-	-	(<-,1]	(<-,1]	(1,->)
Seas+	-	-	-	-	-	{1,2,4,5}	{3,6}	-
Bus	.0200	.0000	.4231	.3158	.2391	.8966	.5714	.4737
Train	.3600	.0000	.5000	.0526	.1087	.1034	.4286	.5263
Other	.6200	1.000	.0769	.6316	.6522	.0000	.0000	.0000
N	50	18	26	19	46	58	21	19
Rule#	10	11	12	13	14	15	16	17

Table 11.4 Continued - alternative transport mode

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Acc	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
D_A	{2,10}	{3,11}	{4}	{4}	{5}	{5}	{5}	{6,9}	{7}
WrkW	-	-	{0,1,2,3,4}	{9}	-	-	-	-	-
Sclss	-	-	-	-	(<-,2]	(2,3]	(3,->)	-	-
Bus	.9615	.4828	.3810	.8148	.8065	.4375	1.000	.6897	.6452
Train	.0385	.4483	.2857	.1481	.1935	.3475	.0000	.3103	.0484
Other	.0000	.0689	.3333	.0371	.0000	.2150	.0000	.0000	.3064
N	78	29	21	27	31	16	16	29	62
Rule#	18	19	20	21	22	23	24	25	26

Finally, for *foreign hotel-based holidays* (R15-27), the train and the "other" alternative modes only seriously threaten the popularity of the bus when the destination is France (when the tourist has already made at least one car-based holiday (R17), or when this is not the case, when the holiday is scheduled for the first part of the summer holidays or for the autumn (R16)), Belgium and Luxembourg or other parts of the world outside Europe (R19), Germany (when there are no missing values on the condition variable "work during the weekend"; R20), or Austria and Switzerland (when the tourist belongs to the third-highest social class; R23).

Table 11.4 Continued - alternative transport mode

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}
Acc	{1}	{2,3}	{4}	{4}	{4}	{4}
D_A	{8}	-	{1,2,5,6,8,9}	{3,4,10,11}	{3,4,10,11}	{7}
Prio4	-	-	-	(<-,1]	(1,->)	-
Bus	.9070	.1667	.7222	.0000	.4762	.1429
Train	.0930	.2222	.1944	.4667	.1429	.0000
Other	.0000	.6111	.0834	.5333	.3809	.8571
N	43	18	36	15	21	21
Rule#	27	28	29	30	31	32

11.3.5 Statistics and validation

To complete this section on tourist transport mode choices, the most important statistics and the predictive abilities of the generated DTs are discussed. The confusion matrix and statistics for the DT for the choice between cars, aeroplanes and alternative means of transport ("main mode of transport") are presented in Table 11.5. Using a deterministic assignment rule, this DT would correctly classify 86.31% of all observations. This means only a small improvement compared to the null-model that would correctly predict 74.92% of the observations, but this is due to the fact that "car" is a very dominant alternative. For the probabilistic DT, the improvement is more important because the number of correctly predicted observations increases from 59.28% for the null-model to 80.21% for the 100 decision rules (35.32% improvement). From the confusion matrix it can be learned that the prediction of air-travel improves enormously, from 12.19% when a probabilistic one-rule model is used to 76.51% using the 100 decision rules. This improved ability of the model to predict these mode choices is probably due to the preferences for aeroplanes to particular destinations. The improvements for the other two mode alternatives are more modest but still noticeable. These conclusions are confirmed by the value of the contingency coefficient, that amounts to 0.7075 out of a maximum of 0.8165 for a 3×100 DT.

Table 11.5 Confusion matrix & statistics for the DT *main mode of transport*

	Car'	Air'	Alter'	Total'
Car	**.8860**	.0183	.0956	.7492
Air	.1065	**.7651**	.1284	.1219
Alter	.5280	.1216	**.3500**	.1289
Total	.7492	.1219	.1289	**.8021**

N° of observations	7121	
Stopping criteria	60 before/25 after	
α	0.05	
N° of columns (= rules)	100	
Theta (1 col)	5335 (.7492)[a]	4221 (.5928)[b]
Theta (100 cols)	6146 (.8631)[a]	5712 (.8021)[b]
χ^2_{DT}(df) / Cont. coef.	7135.53 (198) / .7075	

[a] Deterministic [b] Probabilistic

For the 918 tourist trips that use alternative modes of transport, Table 11.6 shows the statistics and confusion matrix for the choice between the bus-, the train- and the "other"-choice alternatives. Based on the contingency coefficient that amounts to 0.6785 out of a maximum of 0.8165 for a 3×32 DT, it can be concluded that the relationship between the conditions and the responses is reasonable, but slightly less that that of the DT for the choice of the main transport modes. Compared to the null-model, the 32 probabilistic decision rules increase the expected number of correctly predicted observations by 79.45% (from 326 to 585). This improvement is particularly strong for the train- and the "other"-modes, where the numbers of correctly predicted observations increase from 28.32% to 50.55% and from 26.26% to 60.61% respectively. For the bus-mode, finally, the DT improves the predictive accuracy with 62.51% (from 45.32% to 73.65%)

Table 11.6 Confusion matrix & statistics for the DT *alternative transport modes*

	Bus'	Train'	Other'	Total'
Bus	**.7365**	.1621	.1014	.4532
Train	.2617	**.5055**	.2327	.2832
Other	.1718	.2221	**.6061**	.2636
Total	.4532	.2832	.2636	**.6373**

N° of observations	918	
Stopping criteria	35 before/15 after	
α	0.05	
N° of columns (= rules)	32	
Theta (1 col)	416 (.4532)[a]	326 (.3551)[b]
Theta (32 cols)	672 (.7320)[a]	585 (.6373)[b]
χ^2_{DT}(df) / Cont. coef.	783.10 (62) / .6785	

[a] Deterministic [b] Probabilistic

11.4 The question of how much to spend

11.4.1 The choice set

Tourist expenditures for holidays constitute a continuous decision process within which the tourist allocates a certain amount of money to the holiday under consideration. In principle, a "CHAID-based" approach can be used, where an F-test is used to identify groups of observations that differ significantly with regard to the dependent variable (SPSS, 1998). (Strictly speaking this is not CHAID because the χ^2-significance test is not used). Using this approach, for each segment of holidays complying with particular conditions, the average expenditures and standard deviation are be obtained. The disadvantage of this approach is, that in the simulation process negative expenditures can be predicted when the standard deviation approaches or exceeds the mean expenditures. More fundamentally, model assumptions are violated because the F-test requires a normally distributed dependent variable. Alternatively, a regression model for truncated normal distributions, a so-called left-censored Tobin's probit or *Tobit-model*, can be used. Also, expenditures can be divided into ranges, as a result of which a rule-based approach using the CHAID-based algorithm for categorised decision variables is appropriate. *MERCIN* adopts the latter approach because a Tobit-approach systematically overestimated the mean expenditures per person per holiday[36].

[36] Many sets of independent variables were tested, including (combinations of) characteristics of the holiday, the holidaymaker and interactions between these conditions. The lowest predicted mean expenditures per person per holiday approximated NLG 795 while the observed mean amounts to NLG 713.03.

Table 11.7 Choice set for tourist expenditures per person per holiday

label	range	linear mean	observed mean	frequency
P1	<= NLG 55	NLG 27.50	NLG 34.15	730 (10.3%)
P2	NLG 55-120	NLG 87.50	NLG 85.25	712 (10.0%)
P3	NLG 120-187.50	NLG 153.75	NLG 150.59	708 (9.9%)
P4	NLG 187.50-266.53	NLG 227.02	NLG 228.74	698 (9.8%)
P5	NLG 266.53-375	NLG 320.77	NLG 325.23	754 (10.6%)
P6	NLG 375-500	NLG 437.50	NLG 452.80	721 (10.1%)
P7	NLG 500-750	NLG 625	NLG 644.71	722 (10.1%)
P8	NLG 750-1100	NLG 925	NLG 931.24	653 (9.2%)
P9	NLG 1100-1750	NLG 1425	NLG 1400.84	747 (10.5%)
P10	> NLG 1750	NLG 3000[a]	NLG 3008.72	676 (9.5%)
Total		NLG 713.03		7121 (100%)

[a] In the simulation, the upper limit of this segment is set at NLG 4750

In order to obtain categories of expenditures for a rule-based approach, the *expenditures per person per holiday* are categorised into ten segments, i.e. ten groups with approximately equal numbers of observations. The categories, including the number of observations per category are presented in Table 11.7.

MERCIN assumes a linear distribution between the lower and upper limit of each category. Table 11.7 therefore also compares the linear and the observed mean of each segment. For the tenth segment (P10: > NLG 1750), the *observed* upper limit is NLG 14,640. However, in the simulation process, the upper limit is set at NLG 4750 to ensure that the predicted mean matches the observed mean (only 72 of the 676 (11.7%) observed holidays in this segment are more expensive than NLG 4750 per person). Comparing the observed and the linear mean of each segment shows that, except perhaps for the first segment, the linear approximation more or less reproduces the actual mean for each segment of observations.

11.4.2 The condition variables

As before, the condition variables common to all choice facets (Table A1.1 in Appendix 1) and decisions that have been taken previously (Table A1.2) are entered into the CHAID-based algorithm. Only one condition variable is added to these general conditions, i.e. the total amount of costs that have already been planned for more important trips (Table A1.19).

11.4.3 The stopping criteria of the CHAID-based algorithm

For the choice of expenditures too, the α-level for predictor eligibility is set at 5%. The sensitivity analysis is reported in Table A2.17 (Appendix 2). Based on these analysis, the stopping criteria for the choice of the segment of expenditures are set at 120 observations before and 55 after division because the cross-validation measure is only slightly smaller than those of the other models, while the performance at the aggregate level is comparable.

11.4.4 The generated decision table

The generated DT (Table 11.8) for the choice of the segment of expenditures includes 79 decision rules. In this DT, the segments are indicated by their lower limits only to reduce the size of the table. Also, the limits of the fourth and fifth segment are rounded off. In the discussion, the 'actions' are indicated by the lower- and upper-limits of the segments of that contain approximately 10% or more of the observations. The majority of expenditures on foreign short breaks, for instance, amount to NLG 120 (lower limit of the 3rd segment) to 750 (the upper limit of the 7th segment) when the short breaks is the second most important holiday in the annual trip program of a tourist from a household without children (R56).

Table 11.8 DT for the choice of the segment of expenditures

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Fracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Lngth	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,2]	(1,2]	(1,2]
Acc	{1}	{1}	{1}	{2,3,4}	{2,3,4}	{2,3,4}	{2,3,4}	{1}	{2,3,4}	{2,3,4}
Party	{1,2}	{1,2}	{3,4, 5,99}	{1,2}	{3,4, 5,99}	{3,4, 5,99}	{3,4, 5,99}	-	{1,2, 99}	{3,4,5}
#Tp	(<-,1]	(1,->)	-	-	-	-	-	-	-	-
Chld	-	-	-	-	{1}	{2,3}	{2,3}	-	-	-
Seas+	-	-	-	-	-	{1,6}	{2,3,4,5}	-	-	-
DeptD	-	-	-	-	-	-	-	-	{0}	{0}
0	.0080	.0000	.0476	.0714	.0862	.0893	.2564	.0000	.0128	.0140
55	.0800	.0000	.1270	.1071	.2069	.2679	.3162	.0097	.0513	.1399
120	.0960	.0714	.2540	.2143	.2759	.2143	.2564	.0194	.1026	.2727
188	.2560	.1429	.1270	.2857	.2414	.2857	.1111	.0485	.2051	.2517
267	.3120	.3393	.2698	.1667	.1897	.0714	.0427	.1748	.2949	.2308
375	.1520	.2679	.1746	.1310	.0000	.0714	.0085	.2233	.1795	.0629
500	.0720	.1607	.0000	.0119	.0000	.0000	.0085	.2816	.0897	.0070
750	.0240	.0179	.0000	.0000	.0000	.0000	.0000	.1845	.0385	.0210
1100	.0000	.0000	.0000	.0119	.0000	.0000	.0000	.0583	.0256	.0000
1750	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0000	.0000	.0000
N	125	56	63	84	58	56	117	103	78	143
Rule#	1	2	3	4	5	6	7	8	9	10

As expected, the most important condition in the choice of the segment of expenditures is the difference between domestic and foreign trips. This section first discusses the conditions determining the (average) expenditures on domestic holidays (R1-44). Next, expenditures on foreign trips are presented and described (R45-79).

Table 11.8 Continued - segment of expenditures

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Fracc	{0}	{0}	{0}	{0}	{0}	{1}	{1}	{1}	{1}	{1}
Lngth	(1,2]	(1,2]	(1,2]	(2,->)	(2,->)	(<-,2]	(<-,2]	(2,3]	(3,->)	(<-,1]
Acc	{2,3,4}	{2,3,4}	{2,3,4}	-	-	-	-	-	-	-
Party	{1,2,4,99}	{3,5}	-	-	-	-	-	-	-	{1,2}
DeptD	{1}	{1}	{1}	{0}	{1}	-	-	-	-	-
Prio4	(<-,1]	(<-,1]	(1,->)	-	-	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,3]
Work	-	-	-	-	-	{0,2}	{1,3}	-	-	-
0	.0000	.0000	.0330	.0536	.0055	.0339	.0345	.0175	.0000	.0574
55	.0000	.0469	.0659	.0000	.0276	.2034	.0690	.0263	.0098	.1066
120	.0303	.0781	.1538	.1071	.0331	.3051	.2529	.0877	.0784	.2131
188	.1364	.2344	.2088	.1071	.0718	.1695	.1839	.0702	.0098	.2705
267	.2121	.2188	.2418	.1607	.1271	.1695	.1839	.1491	.1667	.2213
375	.3182	.3281	.1429	.2143	.1878	.0678	.1379	.3509	.2451	.1066
500	.2121	.0781	.1209	.2321	.2155	.0339	.1149	.2456	.2843	.0246
750	.0909	.0156	.0220	.0893	.2320	.0000	.0230	.0351	.1176	.0000
1100	.0000	.0000	.0110	.0179	.0718	.0000	.0000	.0088	.0784	.0000
1750	.0000	.0000	.0000	.0179	.0278	.0169	.0000	.0088	.0099	.0000
N	66	64	91	56	181	59	87	114	102	122
Rule#	11	12	13	14	15	16	17	18	19	20

In the case of *domestic holidays*, the possession of a tourist accommodation with a permantent location and the possession of a tourist accommodation without such a location is important. First, there are domestic holidays by *people who do not own any form of accommodation* (R1-15). Under these conditions, hotel-based *short breaks* often cost between about NLG 188 and 500 per person (R1-3), while other short breaks are less expensive at NLG 55-267 (R4-7). Similarly, expenses for *medium long* hotel-based *holidays* amount to NLG 267 to 1100 (R8), while medium long holidays to other accommodation types require NLG 188 to 500 (R9-13). *Holidays of at least nine days*, finally, often cost between NLG 267 and 750 (R14+15), where holidays departing on Friday or Saturday are often at the more expensive end of this range.

Expenditures on domestic holidays by *people who only possess a tourist accommodation without a permanent location* are described by the rules 16 through 31 inclusive. Under these conditions, the priority of the holiday within the annual trip program is the most important condition. When the *most important holiday* is a short break or a medium long holiday, the costs amount to NLG 120 to 374 (R16+17); more extensive trips are also more expensive (267-750; R18+19). For the *second and third most important holidays* in the trip program, the travel party is important. Single travellers and parties of adults only, on average spend the most on their domestic holiday: between NLG 120 and 500 per person (R20+21). Parties with schoolgoing children and/or at least nine people, on the other hand, spend the least, i.e. between NLG 55 and 267 (R22-25). Unknown parties and parties with other children spend between NLG 55 and 375 (R26+27) when their domestic holiday is the second or third most important trip. *Less important holidays*, finally, are also less costly: less than NLG 188 per person (R28-31), except when this holiday is spent in a hotel (NLG 120-750; R28).

Table 11.8 Continued - segment of expenditures

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Peracc	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}
Fracc	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Lngth	(1,->)	(<-,1]	(<-,1]	(1,->)	(1,->)	-	-	-	-	-
Acc	-	-	-	-	-	-	-	{1}	{2,3,4}	{2,3,4}
Party	{1,2}	{3,5}	{3,5}	{3,5}	{3,5}	{4,99}	{4,99}	-	{1,2,4,5}	{1,2,4,5}
Chld	-	-	-	-	-	{1,2}	{3}	-	-	-
DeptD	-	-	-	{0}	{1}	-	-	-	-	-
Prio4	(1,3]	(1,3]	(1,3]	(1,3]	(1,3]	(1,3]	(1,3]	(3,->)	(3,->)	(3,->)
#DT	-	(<-,4]+{99}	(4,->)\{99}	-	-	-	-	-	-	-
Edu	-	-	-	-	-	-	-	-	(<-,5]	(5,->)
0	.0000	.0833	.1895	.0519	.0123	.0179	.1642	.0328	.1781	.0870
55	.0690	.3056	.4211	.2987	.1605	.1607	.2239	.0328	.4247	.2464
120	.1379	.2083	.2212	.2857	.2222	.2321	.1940	.1148	.2466	.2754
188	.2931	.2361	.1053	.2208	.2593	.2143	.2090	.3115	.0822	.1594
267	.2069	.1111	.0421	.0779	.1358	.2143	.1045	.2295	.0411	.1739
375	.1034	.0417	.0212	.0390	.1111	.0536	.1045	.1475	.0137	.0435
500	.1034	.0139	.0000	.0130	.0864	.0893	.0000	.1311	.0137	.0145
750	.0862	.0000	.0000	.0130	.0123	.0179	.0000	.0000	.0000	.0000
1100	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1750	.0001	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
N	116	72	95	77	81	56	67	61	73	69
Rule#	21	22	23	24	25	26	27	28	29	30

Table 11.8 Continued - segment of expenditures

D_NLA	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Peracc	{0}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Fracc	{1}	-	-	-	-	-	-	-	-	-
Lngth	-	(<-,1]	(1,->)	-	(<-,1]	(1,->)	-	-	-	-
Acc	{2,3,4}	{1,2,4}	{1,2,4}	{3}	{3}	{3}	{3}	{3}	{3}	{3}
Party	{99,3}	-	-	-	-	-	-	-	-	-
Chld	-	-	-	-	-	-	{1}	{1}	{1}	{1}
Prio4	(3,->)	-	-	(<-,1]	(1,3]	(1,3]	(3,->)	(3,->)	(3,->)	(3,->)
Edu	-	-	-	-	-	-	(<-,2]	(2,->)	-	-
Inc	-	-	-	-	-	-	(<-,4]	(<-,4]	(4,5]	(5,7]
0	.4300	.0615	.0000	.0370	.6269	.1698	.3103	.5586	.3143	.6344
55	.3300	.1846	.0946	.0833	.2687	.2925	.3966	.2966	.3000	.2903
120	.1900	.2462	.1486	.1204	.0448	.1698	.0862	.0759	.3000	.0323
188	.0300	.1538	.1757	.0370	.0149	.1509	.0862	.0414	.0429	.0323
267	.0100	.1231	.2027	.1574	.0299	.0849	.0862	.0069	.0143	.0000
375	.0100	.1385	.2162	.1759	.0000	.0660	.0345	.0138	.0143	.0108
500	.0000	.0615	.1081	.1759	.0000	.0283	.0000	.0069	.0000	.0000
750	.0000	.0308	.0541	.1019	.0000	.0189	.0000	.0000	.0143	.0000
1100	.0000	.0000	.0000	.0926	.0149	.0000	.0000	.0000	.0000	.0000
1750	.0000	.0000	.0000	.0186	.0000	.0189	.0000	.0000	.0000	.0000
N	100	65	74	108	67	106	58	145	70	93
Rule#	31	32	33	34	35	36	37	38	39	40

With regard to expenditures, the final group of domestic holidays constitutes the trips by *people who own a tourist accommodation with a permanent location* (R32-44). When the holiday is not spent in the own accommodation, these people spend about NLG 55 to 500 on their domestic holiday (R32-33). When the holiday *is spent in the own accommodation*, the holidays are cheaper, except for when this is the most important holiday in the annual tourist trip program (NLG 120-1100; R34). When the second most important holiday is a short break, people on average spend less than NLG 120 (R35), while under these conditions holidays of at least 5 days cost between NLG 0 and 267 (R36). When less important domestic holidays by people with a permanent accommodation are concerned, tourists without children spend slightly more per person (NLG 0-188; R37-41) than tourists from households with children (NLG 0-120; R42-44).

The second part of the DT is concerned with expenditures on *foreign holidays* (R45-79). For foreign holidays, the presence of children in the household is important, where people without children on average spend more per person than people with children in the household. When the *most important holiday* of *people without children* is a short break or a medium long holiday by car or train, the costs of the foreign holiday vary between NLG 188 and 1100 per person (R45). Longer holidays by car or train, and holidays by bus (regardless of the duration) are most likely to cost at least NLG 500 (R46-49+55). The most expensive trips, however, are most important holidays abroad by aeroplane that do not sell below NLG 1100 per person (R50-54).

Table 11.8 Continued - segment of expenditures

D_NLA	{1}	{1}	{1}	{1}	{2}	{2}	{2}	{2}	{2}	{2}
Peracc	{1}	{1}	{1}	{1}	-	-	-	-	-	-
Lngth	-	-	-	-	(<-,2]	(2,3]	(2,3]	(3,->)	(3,->)	(<-,3]
Acc	{3}	{3}	{3}	{3}	-	{1,3}	{2,4}	-	-	-
Chld	{1}	{2,3}	{2,3}	{2,3}	{1}	{1}	{1}	{1}	{1}	{1}
Prio4	(3,->)	(3,->)	(3,->)	(3,->)	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]
#DT	-	(<-,0]	(0,->)+{99}	(0,->)+{99}	-	-	-	-	-	-
Inc	(7,->)	-	-	-	-	-	-	-	-	(<-,4]
Sclss	-	-	(<-,3]	(3,->)	-	-	-	-	-	-
Tmode	-	-	-	-	{1,4}	{1,4}	{1,4}	{1,4}	{1,4}	{2}
Region	-	-	-	-	-	-	-	{1,2}	{3}	-
0	.2113	.9589	.6697	.8306	.0000	.0000	.0000	.0000	.0000	.0000
55	.2254	.0411	.2294	.1371	.0217	.0167	.0000	.0000	.0000	.0093
120	.3662	.0000	.0367	.0323	.0543	.0000	.0000	.0000	.0000	.0000
188	.0986	.0000	.0550	.0000	.1196	.0167	.0122	.0000	.0132	.0093
267	.0704	.0000	.0092	.0000	.1413	.0333	.0610	.0000	.0132	.0000
375	.0282	.0000	.0000	.0000	.2609	.0000	.0976	.0195	.0658	.0370
500	.0000	.0000	.0000	.0000	.1413	.1500	.2439	.0714	.1053	.0370
750	.0000	.0000	.0000	.0000	.1957	.1167	.2317	.1169	.1842	.0926
1100	.0000	.0000	.0000	.0000	.0435	.4667	.2561	.4026	.2895	.3611
1750	.0000	.0000	.0000	.0000	.0217	.1999	.0975	.3896	.3288	.4537
N	71	73	109	123	92	60	82	154	76	108
Rule#	41	42	43	44	45	46	47	48	49	50

For the *second most important holiday* by people without children, again the duration and the transport mode condition the expenditures. As expected, under these conditions, people spend less on short breaks (NLG 120 to 750; R56). Medium long holidays by aeroplane cost at least NLG 750 (R59). When other means of transport are used, medium long holidays to Belgium and Luxembourg and Germany require NLG 188-750 (in addition, 23.33% of these trips belong to the most expensive segment (R58)), while other destinations abroad amount to NLG 375 to 1750 (R57). Holidays of at least 9 days, finally, cost at least NLG 1100 when the destination is reached by aeroplane (R62), while trips with other means of transport require at least NLG 500 per person (R60+61).

For *less important holidays* by people without children, too, duration and travel mode are important. Again, trips by air are the most expensive (at least NLG 750; R67), while people spend the least amount of money on short breaks (NLG 55 to 500; R63-64). Holidays of at least 5 days by car or “other” transport modes to Belgium, Luxembourg, Germany, Scandinavia and Denmark cost between NLG 267 and 1100 (R66), while expenditures on holidays with these characteristics to other destinations and other foreign holidays by bus or train on mainly amount to NLG 375 to 1750 (R65+68).

Table 11.8 Continued - segment of expenditures

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Fracc	-	-	-	-	-	-	-	-	-	{0}
Lngth	(<-,3]	(<-,3]	(3,->)	(3,->)	-	(<-,1]	(1,2]	(1,2]	(1,2]	(2,->)
Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}
Prio4	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,2]	(1,2]	(1,2]	(1,2]	(1,2]
Work	-	-	{0,1}	{2,3}	-	-	-	-	-	-
Inc	(4,->)	(4,->)	-	-	-	-	-	-	-	-
Tmode	{2}	{2}	{2}	{2}	{3,5}	-	{1,3,4,5}	{1,3,4,5}	{2}	{1,3,4,5}
WrkW	(<-,2]	(2,->)+{9}	-	-	-	-	-	-	-	-
D_A	-	-	-	-	-	-	{1,2,5,6,7, 8,9,10,11}	{3,4}	-	-
0	.0000	.0000	.0000	.0000	.0000	.0225	.0000	.0000	.0000	.0000
55	.0000	.0000	.0000	.0000	.0000	.0337	.0135	.0167	.0000	.0000
120	.0000	.0000	.0000	.0000	.0068	.1124	.0000	.0167	.0000	.0000
188	.0000	.0000	.0000	.0000	.0068	.1910	.0541	.1167	.0000	.0000
267	.0000	.0000	.0000	.0000	.0000	.1685	.0541	.2167	.0000	.0119
375	.0000	.0000	.0000	.0143	.0411	.1685	.1216	.1833	.0000	.0357
500	.0000	.0149	.0000	.0143	.1027	.2022	.1216	.1833	.0492	.0833
750	.0364	.1045	.0000	.1000	.2260	.0562	.3378	.0333	.2459	.2262
1100	.1818	.3433	.0459	.0000	.3630	.0449	.2297	.0000	.4426	.4405
1750	.7818	.5373	.9541	.8714	.2536	.0000	.0676	.2333	.2623	.2024
N	55	67	109	70	146	89	74	60	61	84
Rule#	51	52	53	54	55	56	57	58	59	60

The final part of the DT for expenditures is concerned with *foreign holidays of people with children* in the household. When these people spend their foreign holiday in a *hotel*, the costs usually amount to NLG 375 to 1750 when these people travel by bus or train (R72), while the costs often exceed NLG 750 when people travel by aeroplane (R71). When people travel by other means (including car), they spend NLG 267 to 750 on holidays to France, Belgium and Luxembourg, Germany, Scandinavia and “other destinations outside Europe” (R69), while the costs of holidays to all other destinations amount to NLG 500 to 1750 (R70).

Table 11.8 Continued - segment of expenditures

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Fracc	{1}	-	-	-	-	-	-	-	-	-
Lngth	(2,->)	(2,->)	(<-,1]	(<-,1]	(1,->)	(1,->)	-	-	-	-
Acc	-	-	{1}	{2,3,4}	-	-	-	-	{1}	{1}
Chld	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{1}	{2,3}	{2,3}
Prio4	(1,2]	(1,2]	(2,->)	(2,->)	(2,->)	(2,->)	(2,->)	(2,->)	-	-
Tmode	{1,3,4,5}	{2}	{1,5}	{1,5}	{1,5}	{1,5}	{2}	{3,4}	{1,5}	{1,5}
D_A	-	-	-	-	{1,2,5,6, 7,8,9,11}	{3,4,10}	-	-	{1,3,4, 10,11}	{2,5,6,7, 8,9}
0	.0000	.0000	.0238	.0833	.0145	.0000	.0000	.0123	.0504	.0000
55	.0143	.0000	.0119	.2167	.0145	.0333	.0000	.0247	.0588	.0000
120	.0000	.0000	.1429	.1833	.0000	.1000	.0117	.0247	.0924	.0000
188	.0286	.0000	.2381	.1667	.0000	.0333	.0000	.0741	.0840	.0230
267	.0571	.0119	.2976	.2000	.0870	.2667	.0177	.0864	.2017	.0230
375	.0286	.0119	.2143	.1000	.1594	.2333	.0708	.1358	.1597	.0805
500	.2143	.0119	.0595	.0333	.2029	.1500	.0619	.1728	.1849	.2414
750	.2000	.0476	.0000	.0167	.2464	.1333	.1858	.2593	.0840	.2989
1100	.3286	.2500	.0119	.0000	.2029	.0333	.4159	.1975	.0756	.2989
1750	.1285	.6667	.0000	.0000	.0724	.0168	.2362	.0124	.0085	.0343
N	70	84	84	60	69	60	113	81	119	87
Rule#	61	62	63	64	65	66	67	68	69	70

Table 11.8 Continued - segment of expenditures

D_NLA	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}	{2}
Lngth	-	-	(<-,3]	(3,->)	-	-	-	-	-
Acc	{1}	{1}	{2,3,4}	{2,3,4}	{2,3,4}	{2,3,4}	{2,3,4}	{2,3,4}	{2,3,4}
Chld	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}	{2,3}
Prio4	-	-	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(<-,1]	(1,2]	(2,->)
Tmode	{2}	{3,4}	-	-	-	-	-	-	-
D_A	-	-	{1,5,7,8,10}	{1,5,7,8,10}	{1,5,7,8,10}	{2,6,9,11}	{3,4}	-	-
HHsz	-	-	(<-,4]	(<-,4]	(4,->)	-	-	-	-
0	.0063	.0090	.0000	.0000	.0187	.0112	.0074	.0140	.1650
55	.0000	.0180	.0000	.0000	.0093	.0000	.0370	.1049	.2233
120	.0127	.0270	.0164	.0000	.0000	.0000	.0667	.1399	.1845
188	.0000	.0360	.0000	.0000	.0187	.0112	.0963	.1888	.1456
267	.0000	.0721	.0656	.0138	.0748	.0112	.1333	.1985	.1262
375	.0253	.1171	.1311	.0483	.1963	.0225	.2519	.1608	.0874
500	.0443	.2523	.2787	.2621	.3925	.1348	.2815	.1049	.0291
750	.1582	.1892	.2951	.4138	.1963	.2360	.1037	.0629	.0194
1100	.3797	.2523	.1967	.2069	.0935	.3933	.0222	.0210	.0194
1750	.3735	.0270	.0164	.0551	.0000	.1798	.0000	.0043	.0000
N	158	111	61	145	107	89	135	143	103
Rule#	71	72	73	74	75	76	77	78	79

Foreign holidays of people with children in the household that are *not spent in a hotel*, finally, are less expensive, especially when the trip is the second most important (NLG 55 to 750; R78) or even less important (NLG 0 to 375; R79) holiday in the annual trip program. When the non-hotel-based foreign holiday is the most important holiday of that year, the costs amounts to NLG 267 to 1100 when the destination is Germany, Austria or Switzerland (R77), to at least NLG 500 when the destination is Spain, Portugal, Italy, Greece, the South East Mediterranean area or "other" parts of the world outside Europe (R76), and to NLG 375 to 1750 when other destinations are concerned (R73-75).

11.4.5 Statistics and validation

The statistics and the confusion matrix for this final DT of the MERCIN-system are presented in Table 11.9. In order to reduce the size of this table, the expected fraction of correctly predicted cases is presented in three decimals only, while the segments are labelled from P1 through P10 instead of by the lower and/or upper limits (see also Table 11.7).

Table 11.9 Confusion matrix & statistics for the DT *segments of expenditures*

	P1'	P2'	P3'	P4'	P5'	P6'	P7'	P8'	P9'	P10'	Total'
P1	**.508**	.226	.108	.064	.040	.027	.014	.007	.005	.002	.103
P2	.232	**.240**	.177	.130	.097	.062	.033	.017	.009	.004	.100
P3	.111	.178	**.196**	.160	.146	.100	.061	.028	.013	.006	.099
P4	.067	.133	.164	**.189**	.173	.126	.081	.041	.018	.008	.098
P5	.039	.092	.136	.160	**.188**	.155	.118	.064	.034	.014	.106
P6	.027	.061	.097	.122	.162	**.179**	.157	.100	.065	.030	.101
P7	.014	.033	.059	.078	.123	.156	**.198**	.155	.129	.056	.101
P8	.007	.019	.030	.045	.077	.113	.174	**.205**	.212	.119	.092
P9	.005	.008	.012	.018	.035	.064	.125	.186	**.306**	.241	.105
P10	.002	.003	.004	.006	.010	.028	.055	.118	.256	**.517**	.095
Total	.103	.100	.099	.098	.106	.101	.101	.092	.105	.095	**.254**

N° of observations	7121	
Stopping criteria	120 before/55 after	
α	0.05	
N° of columns (= rules)	79	
Theta (1 col)	754 (.1059)[a]	713 (.100)[b]
Theta (79 cols)	2606 (.3660)[a]	1806 (.2536)[b]
χ^2_{DT}(df) / Cont. coef.	12,290.95 (702) / .7957	

[a] Deterministic [b] Probabilistic

Using the 79 probabilistic decision rules, the DT is able to correctly predict 25.36% of the observations. This may appear to be a very poor performance compared to the previously discussed DTs, but compared to the (probabilistic) null model this presents an improvement of more than 150%. This is due to the fact that the choice set consists of 10 segments that, by definition, contain approximately 10% of the observations. Looking at the ability to predict the separate segments of expenditures reveals that the DT is especially adapted to the segments at both ends of the choice sets (P1, P9 and P10), while the segments in the middle of the scale still improve by 77% to 140%. Furthermore, the confusion matrix shows that the majority of "confusion" takes place between adjacent segments and not between remote segments. In other words, cheap holidays are very unlikely to be classified as very expensive and the other way around. In this respect, the DT performs very well. The contingency value also indicates the DT to perform very well: 0.7957 on a maximum of 0.9487 for a 10×79 DT.

11.5 Conclusion and discussion

This chapter discussed the final DTs of the *MERLIN*-system, including the DTs for the choices of accommodation, transport mode and expenditures. The accommodation alternatives considered by *MERLIN* are predominantly based on ownership. With regard to accommodations that are owned by the tourist, the choice alternatives include facilities with a permanent location, and commodities such as boats, tents and caravans without a permanent location. Lodging facilities that are not owned by the tourist are further divided in "hotels, motels, pensions, apartments, and rooms without pension" and "other accommodations not owned by the tourist". Given this set of choice alternatives, it was no surprise to find the choice of accommodation to have a strong relation with the possession of an accommodation. In interaction with these conditions, the destination proved to be important, where hotels (etc.) are very popular for destinations abroad. However, for longer holidays (2 weeks or more) abroad, the own accommodation is more important. This is probably due to financial considerations.

The second holiday choice was the choice of transport mode. This choice facet is described using two DTs, the first of which describes the circumstances that condition the choice between the options "car", "aeroplane" and "alternative". Decision rules distinguishing between these alternative, (often) more ecologically sound modes were the subject of the second DT for transport modes. In both DTs the importance of domestic and foreign destination was unambiguous. First and foremost, due to the size of the Netherlands, the "aeroplane" is never selected for domestic holidays. The car is by far the most popular means of transport for domestic trips. Even people without a car select this mode for 66.02% of their domestic holiday!

Although aeroplanes and the alternative mode gain importance for foreign holidays, still 56.25% of all holidays abroad are undertaken by car. Aeroplanes are particularly popular for foreign hotel-based holidays. Also, there is a strong relation with the presence of children in the household, where childless households select the aeroplane more frequently, and particular destinations. The later relation is probably based on distance and/or the offer of package tours to certain destinations. The bus is the prevailing alternative mode for foreign holidays.

The final conclusion with regard to the DT for the choice of transport modes is that the three variables representing the tourist's general propensities for transport modes were not included in the decision rules. Recall that these propensities were calculated using the MNL-model discussed in section 9.2, and included in the DTs for the destination choices to make these choices susceptible to general changes in transport mode preferences. The absence of these conditions in the decision rules for transport modes implies that user-defined changes in the aggregate preferences for these transport modes will only affect mode choice via the effect on destination choices. Since chapter 9 already indicated that shifts in transport mode propensities are not likely to have significant effects on destination choices, this carry-over effect to transport mode choices will probably not come off.

The final DT describes the choice of the segment of expenditures per person per trip. The DT again showed the destination to be very important, where, evidently, holidays to domestic destinations are, on average, cheaper than trips abroad. For domestic holidays, the possession of tourist accommodations proved to be important. People with such accommodation spend significantly less per person than people without the opportunity to use their own facilities: for 77-80% of the domestic trips, people with a permanent accommodation often spend less than NLG 188; for people who only possess a tourist accommodation without a permanent location the expenditures are between NLG 55 and 500, while people without any facilities spend between NLG 120 and 750 per person per trip. Although people who are in possession of a tourist accommodation spend less financial resources per person per holiday, their expenditures on holidays on an annual base may equal or even exceed those made by other tourists due to higher trip frequencies (see chapter 7).

With regard to foreign holidays, the presence of children in the household proved to be most significant because tourists without children tend to spend more (per person per holiday) than other tourists. Finally, in addition to these differences, the transport mode and the duration of the holiday dominate the expenditure choices where longer holidays and holidays by air are much more expensive.

12 Validation and Demonstration

This chapter brings together the conceptual and empirical models that were advanced in the previous chapters. Next, the validity of the system is discussed. Also, three scenarios for the future are simulated and discussed to demonstrate the use and output of the system. Finally, some general conclusions will be drawn with regard to development and functioning of the system.

12.1 Introduction

The previous chapters have successively discussed the conceptual considerations underlying the *MERLIN*-system, the data collection process and the empirical models that were induced based on the obtained data. The present chapter brings together these conceptual and empirical building blocks and discusses the validation and the working of the *MERLIN*.

In terms of the schemes that were introduced in chapter 4, *MERLIN* is a stochastic discrete static and empirically tested microsimulation model that is implemented using a general programming language. The system is *stochastic* because the utility-based models and decision rules comprise random elements. Random numbers are used to determine the outcome of each stage in the tourist choice process. As a consequence, multiple simulation outcomes (= runs) are required, and the simulation results are presented using statistical measures such as the mean value and the standard deviation. Simulation results are compared using *t*-values, i.e. testing the significance of the difference between two simulation outcomes given the standard deviations and the degrees of freedom.

MERLIN is a *discrete* simulation model because the dependent variable, i.e. a tourist trip pattern including trip profiles, is mainly discrete in nature. It is a *microsimulation* model because trip patterns are simulated for each member of the population individually. Aggregate results are obtained by putting together these individual simulations to describe the tourist choices of the entire population.

The present version of *MERLIN* conveniently assumes the 1998 CVO-sample of respondents to represent the population of interest. Since time does not play an essential role and these data are cross-sectional, the system can be characterised as a *static* simulation model. The characteristics of the population are changed by bringing about changes in the composition of this sample and/or the characteristics of its units. Section 12.4 discusses which options are available for changing the population, and how these changes are implemented. In other words, this section discusses how the characteristics of the 1998 CVO-sample are adjusted to produce the population of simulation units.

MERCIN is coded in Delphi, i.e. a *general programming language* because this offered the flexibility required to build such a large and complex system. The user manual, including set-up instructions, can be found in Appendix 4. Finally, *MERCIN* can be characterised as a *model-based simulation system* because both the empirical building blocks and the system as a whole are tested empirically. The rule bases that were discussed in the previous chapters have each been tested empirically using confusion matrices and various statistical measures that represent the ability of the decision rules to reproduce the original data. Also, in determining the stopping criteria of the CHAID-based algorithm, sensitivity analyses were used based on resubstitution and cross-validation measures. With regard to the trip generation/time allocation models, several statistical measures were presented. However, the models were not tested on their ability to reproduce the original input data. Chapters 6 and 7 discussed the discrepancies between the data that were used to calibrate the trip generation/time allocation models and the 1998 CVO-data for adults (see also Table 6.2). Given the importance of this first empirical model within the system, a weighting scheme is introduced to obtain a better match between the predicted trip frequencies based on the trip generation/time allocation models and the CVO-data. Section 12.2 accounts for the induction of these final empirical changes. Subsequently, the validation and working of the entire *MERCIN*-system are discussed in terms of the stability of the simulation results, the reproduction of the original data at the aggregate level, and the correlation between observed and predicted trip frequencies, destination and duration choices and expenditures at the individual level.

This chapter is organised as follows. First, section 12.2 discusses the final adjustments to the empirical building blocks of the system. Subsequently, the validity of the entire system is discussed. Next, section 12.4 lists the available options for specifying projects that analyse the impact of possible scenarios for the future on tourist trips patterns. These options are illustrated in the subsequent section that demonstrates the working of the system by discussing the results of three of these projects. Finally, in section 12.6 some general conclusions are drawn with regard to development and functioning of the system.

12.2 Weighting scheme for the time allocation/trip generation stage

12.2.1 Introduction

In chapter 6, the characteristics of the data that were used to induce the empirical models were discussed. With regard to the data that were used to calibrate the trip generation/time allocation models for adults, special attention was paid to the correspondence between the entire 1998 CVO adult (≥ 16 years) sample and the 1530 "usable allocation responses" that were obtained from the responses to supplementary questions in the CVO-questionnaire. The reason for this testing was that the sequential trip generation/time allocation models (see Chapter 6) were to be based on the "usable allocation responses" while they are used to simulate the choice behaviour of the entire adult 1998 CVO-sample. In chapter 6 it was concluded that there were some notable differences in the trip frequencies in the two (sub)sets. In particular, the average number of short breaks in the "usable allocation responses" was (seriously) overestimated, whereas the average number of extra long holidays was (seriously) underestimated (see Table 6.2). Hence, it was decided to introduce a weighting scheme to obtain a better match between the predicted trip frequencies based on the trip generation/time allocation models and the CVO-data. If, despite these adjustments for data differences, the simulation results would still deviate from the observed participation and trip choices, an additional weighting scheme can be introduced (the predictions based on "child" models do not have to be adjusted for data deviations because they are based on the original CVO-data; however, the need for any type of weighting should also be tested for these models). This section details the induction of these weighting schemes. This also includes a discussion of the way the utility-based time allocation models presented in Chapter 7 are used to generate trip frequencies.

12.2.2 Formula's

The trip generation/time allocation models include two stages - the participation and the trip choices. Weighting schemes should therefore be introduced for each stage individually. In the *participation choice models* that are based on single choice data, the weight ($¥_p$) for each participation choice alternative *p* is included by multiplying the predicted probabilities for selecting a participation choice alternative *P(p)* (where *p* is the participation alternative "day-trips only", "holidays only" or "both day-trips and holidays") by these factors. In formula:

$$P^{adj}(p) = ¥_p * P(p)$$

where $P^{adj}(p)$ is the adjusted probability of selecting participation choice alternative *p* and *P(p)* is the probability of selecting participation choice *p* given the characteristics of the (potential) tourist and based on the model parameters presented in Table 7.4. Next, the probability for the reference alternative ("no tourist trips") is obtained by subtracting the adjusted probabilities for the other alternatives from one. Theoretically, this procedure could result in predicted (cumulative) probabilities that exceed 1 (i.e., *P*("*day-trips only*") + *P(*"*holidays only*") + *P(*"*both day-trips and holidays*") > 1). In practice, however, the predicted probabilities adjusted for the induced weights never cause any problems.

Finally, based on these adjusted probabilities $P^{adj}(p)$ for each participation choice alternative *p*, the participation choice is determined using random numbers.

In the *trip choice models* that are based on frequency data, first, the probabilities for each trip type *P(t)* (where *t* is the trip type, i.e., day-trip, short break, extended holiday, long holiday or extra long holiday) are calculated using the appropriate utility-based model (depending on the participation choice; see Tables 7.5 through to 7.7 inclusive). These probabilities *P(t)* represent the probability that, given the characteristics of the tourist, one day will be allocated to that particular trip type. Multiplying these probabilities *P(t)* by 365 yields the number of days *D(t)* a tourist will allocate to a particular trip type *t* per year. The weight for each trip type Y_t is included by multiplying the allocated number of days *D(t)* by these factors (again, in practice this does not give rise to any problems.). This adjustment factor consists of an adjustment for data imperfections (Ψ_t) and an additional adjustment factor that ensures that the predicted trip frequencies match the observed number of trips (¥$_t$). In formula:

$$D^{adj}(t) = Y_t * D(t) = Y_t * \left(365 * P(t)\right) = \left(\Psi_t * ¥_t\right) * \left(365 * P(t)\right)$$

where $D^{ajd}(t)$ is the adjusted amount of days allocated to trip type *t* and all other symbols as before.

To obtain the trip frequencies *F(t*) for each trip type *t*, the adjusted amount of days allocated to each trip type is divided by the number of days that were used to convert the observed trip frequencies to "the number of days allocated to a tourist trip" C_t (recall that the "conversation days per trip type" were set at: $C_{day\text{-}trip} = 1$; $C_{short\ break} = 3$; $C_{medium\ long\ holiday} = 7$; $C_{extended\ holiday} = 12$; $C_{long\ holiday} = 22$; and $C_{extra\ long\ holiday} = 30$; see chapter 7). These trip frequencies are rounded down to the nearest integer, resulting in a number of trips that will be made "for certain" ($INT(D^{ajd}(t), C_t)$), the number of days associated with these trips ($C_t * INT(D^{ajd}(t), C_t)$), and a remainder of days that should be allocated to trip type *t* ($MOD(D^{ajd}(t),C_t)$). Finally, this remainder of days is divided by the "conversation days per trip type" C_t to obtain the probability that an extra trip of trip type *t* will be made. Using these probabilities and a random number, the choice for extra trips is determined. In formula:

$$F(t) = INT\left(D^{adj}(t), C_t\right) + \lambda, \qquad \lambda \begin{cases} = 1, & if \quad R_{mc} < \left(MOD\left(D^{adj}(t), C_t\right)\right) \\ = 0, & if \quad R_{mc} \geq \left(MOD\left(D^{adj}(t), C_t\right)\right) \end{cases}$$

where:

F(t) is the number of trips of type *t*

$INT(D^{ajd}(t), C_t)$ returns $D^{ajd}(t)$ divided by C_t and rounded down to the nearest integer

$MOD(D^{ajd}(t), C_t)$ returns the remainder when $D^{ajd}(t)$ is divided by C_t

R_{mc} is a random number between zero and one ($0 \leq R_{mc} < 1$)

and all other symbols as before.

12.2.3 The induced weighting schemes

Initially, general weighting schemes were introduced for the entire sample. Using these weights, the first simulation results of the *MERLIN*-system (that is the simulation of the entire tourist decision-making process) indicated a significant underestimation of the accommodation types “personal accommodation without a permanent location” and “permanent personal accommodations”. Further examination of the discrepancies between the “allocation data” (based on the responses to additional questions presented to the adult sample) and the number of holidays according to the 1998 CVO-data revealed significant differences between the segments of the sample owning such types of accommodations and the segments without personal accommodations. Other segmentations of the sample that were also examined, including segmentations based on civil state, household size and income. No significant differences were found for these groups. It was therefore decided to induce different weighting schemes for the following segments:

1. adults in possession of a tourist accommodation with a permanent location (*N* = 215);
2. adults in possession of a tourist accommodation without a permanent location (who do not own a tourist accommodation with a permanent location; *N*= 915);
3. adults without a tourist accommodation (*N* = 1706).

The weighting schemes are based on the simulation results after 1000 runs; the formulas for inducing the various weights are presented in Appendix 3. Subsequently, the weighting schemes for the "adult" are presented in section A3.1 and the tests for the "child" models are presented in section A3.2.

The "child" model perfectly reproduces the observed data (see section A3.2). For the "adult" model, on the other hand, weights are required for the participation model (($¥_p$), and in the trip models for data discripancies (Ψ_t) and unexplained factors ($¥_t$) to obtain a perfect match between the observed data ("98CVO" in Appendix 3) and the predicted number of holidays (and day-trips). The weights for the participation choice model vary between 0.740 and 1.19. For this model, the "day-trips only" alternative is always scaled down (0.740, 0.930 and 0.950), while the "both day-trips and holidays" alternative requires weights equal to or exceeding one in all case. The combined trip frequency weights ($Y_t = \Psi_t * ¥_t$) vary significantly more: between 0.542 and 3.99. As expected (based on Table 6.2), weights for short breaks are always smaller than one (0.963, 0.677 and 0.542) while factors for extra long holidays always exceed one. The extra long holidays require the largerst weights (3.99, 1.53 and 1.86). The weights for the segment with adults in possession of a tourist accommodation with a permanent location deviate the most from one, both for the participation and the trip frequency choices. The weights for the adults in possession of a tourist accommodation without a permanent location, on the other hand, come closest to one. These weighting schemes represent the final input into the MERCIN-system.

12.3 Validation

This section discusses the validation results of the entire system. If the system is capable of successfully reproducing the empirical data, there is at least some evidence of the validity of the assumptions. In the previous chapters and sections, the empirical tenability of the various building blocks proved to be rather satisfactory in most cases. By testing the entire system, however, the assumptions with regard to the relationships between the system's components (including the decomposition of the entire system into these components) are tested. In this chapter, therefore, the performance of a baseline simulation the MERCIN-system is tested at both the individual and the aggregate level.

As argued before (see chapter 4), a *baseline simulation* describes the 'no change' simulation to which other simulation experiments are compared (Merz, 1991). In case of the MERCIN-system, the most important baseline simulation is the default reference project "*MERCIN*-null" that represents the simulated tourist trip patterns of the 3562 members of the 1998 CVO-panel without any form of upgrading and/or ageing. Experiments that use the entire 1998 CVO-panel to compose the population of simulation units should be compared to this reference rather than to the original 1998 CVO-panel data to exclude system irregularities from the impact assessment (alternatively, if only a particular segment of the population is simulated, first a baseline simulation for that segment should be created; next, simulation results of changes to this segment should be compared to this reference; see, for example, section 12.5.1). The aim of this section, then, is to discuss the presence and extent of these system irregularities by comparing the predicted tourist trip patterns for the 1998 CVO-panel to the observed patterns both at the aggregate and the individual level. First, however, the stability of the simulation results of *MERCIN*-null is discussed.

12.3.1 Stability of the simulation results

The *MERCIN*-system is a stochastic microsimulation model, which implies that the output for a given input is uncertain and statistical operations on multiple simulation outcomes are required when conducting simulation experiments. The *first test for the stability* of the simulation outcomes is therefore to view the course of the mean value of the dependent variable for the default reference project "*MERCIN*-null" across multiple simulation runs. The *second test* of stability comprises a statistical comparison between the results of the default reference project "*MERCIN*-null" and an identical simulation experiment[37].

As mentioned before, "*MERCIN*-null" represents the simulated tourist trip patterns of the 3562 members of the 1998 CVO-panel. For this default reference project, the number of runs is set at two-hundred runs (N = 200). Based on conceptual considerations, the dependent variable in the *MERCIN*-system, i.e. tourist trip patterns, is composed of various choice facets, including, the number of day-trips and the number of holidays, and for each holiday the duration, the travel party, the timing, the destination, the accommodation, the mode of transport, the date of departure and the expenditures. And for each simulation experiment, the *MERCIN*-system records the mean value and the standard deviation across the

[37] A third test for stability was conducted by dividing the 1998 CVO population into two equal halves and comparing the simulation results for the two segments. Although the two segments comprised an equal number of members, their compositions (in terms of socio-economic variables) proved to differ slightly, apparently resulting in many significant differences in the simulation results between the two halves. The results of this test are not discussed in detail.

simulation runs for each facet of the dependent variable. These simulation results are presented in the *standard output-file*. Appendix 4 (section A4.3.1) explains the composition and the interpretation of this standard output-file, and Appendix 5 displays the standard output-file that compares the default reference project "MERLIN-null" to an identical test project.

First, for each indicator of the tourist trip pattern (as presented in the standard output file), the course of the mean value for the standard reference "MERLIN-null" was plotted against the number of runs. For the majority (85-90%) of the indicators, the eventual value is reached (± 2.5%) after 25 to 50 simulation runs. All indicators eventually resulted in a stable value.

There was no apparent pattern that could explain why some indicators did not reach their eventual value fast and/or by a steady course (e.g. earlier or later in the tourist choice process, smaller or larger values). These indicators do, however, tend to have a larger standard deviation relative to the mean value (which should not be a surprise). Figure 12.1 shows an example of an indicator that reached the eventual value rather steady while Figure 12.2 shows a less steady course.

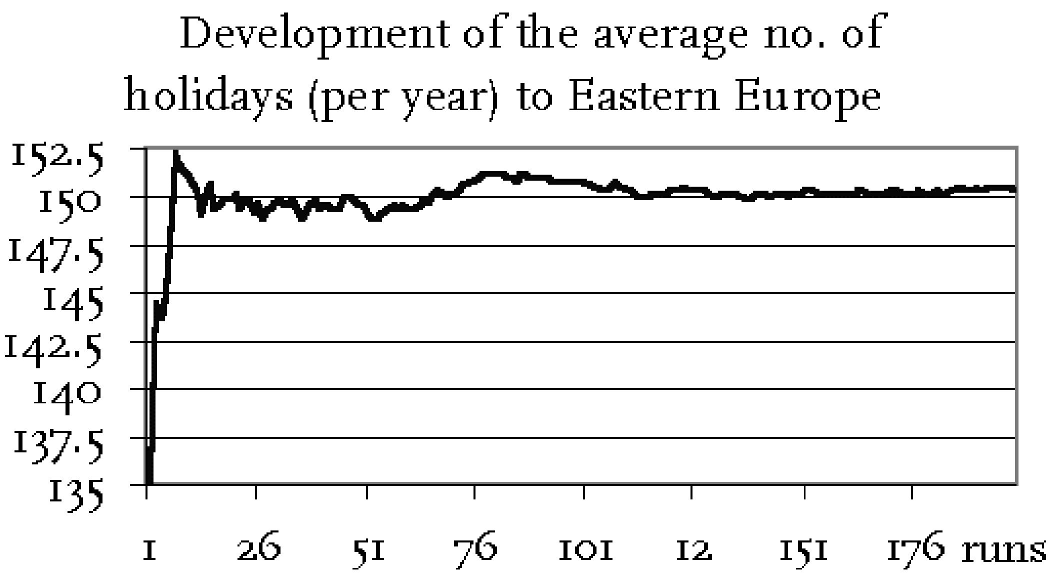


Figure 12.1 A steady course of the mean value of an indicator for the standard reference project "MERLIN-null" plotted against the number of runs

The second test for the stability of the simulation results statistically compared the default reference project "MERLIN-null" to an identical test project. The results are presented in Appendix 5, Tables A5.1 (participation choices) and A5.2 (holiday choices). In this appendix, 106 indicators of the tourist trip patterns are presented and compared across the two identical projects. Out of these

106 indicators, 3 indicator values of the test project differ more than 1% from "*MERLIN*-null" (i.e. the average number of children without tourist trips, the number of holidays to South Each Mediterranean countries and the average expenditures on domestic long holidays), while 2 indicators differ even more than 2% (the average number of extra long holidays per adult unit who only goes on holidays and the average number of extra long holidays per unit younger than 16 years). None of these differences, however, are significant at the 5% level. This is probably due to the inherent variability of these indicators as indicated by the large standard deviation relative to the mean value.

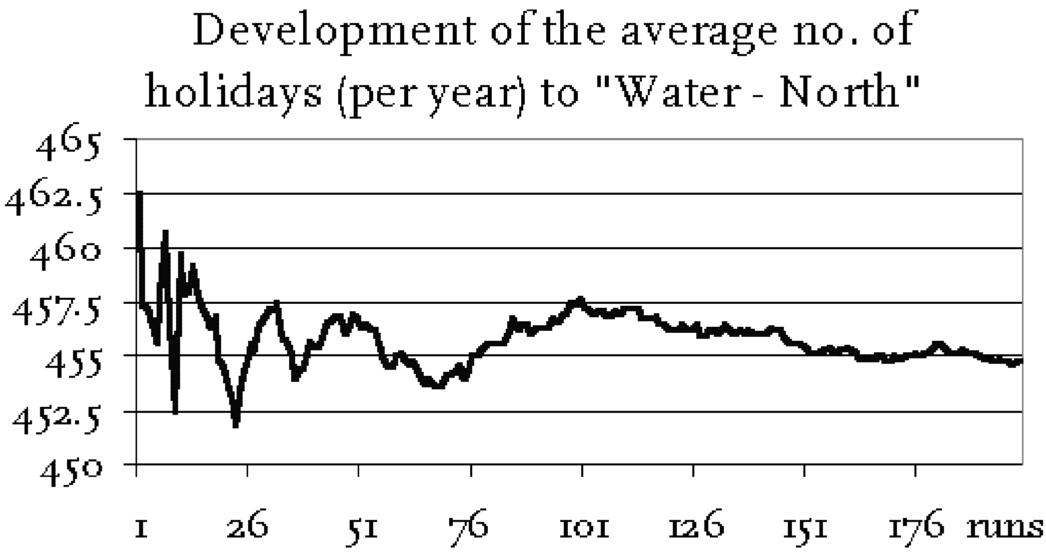


Figure 12.2 A less steady course of the mean value of an indicator for the standard reference project "*MERLIN*-null" plotted against the number of runs

Significant differences at the 5% level are found for 3 indicators (2 of which are related), i.e. the average number of short breaks for children who go on holidays, the average number of holidays departing in the first part of the summer school holiday period and the average number of holidays departing in the final part of the summer school holiday period. These differences, however, are based on chance because when the test is repeated, three to four significant differences are found for different indicators time and again[38].

[38] Both "*MERLIN*-null" and the identical test project are based on 200 simulation runs. Similar tests for stability were also conducted using identical projects based on 25, 50 and 100 runs. In these tests, the number of significant differences between two identical projects appeared to increase slightly when the number of runs was decreased. However, the total number of significant differences never exceeded 5. In addition, the indicators that showed significant differences in these tests never coincided across the various tests.

In summary, therefore, in can be concluded that the stability of the simulation results is satisfactorily both in terms of the convergence of the mean value across consecutive simulation runs and the reproductive power for identical input conditions. The homogeneity of the population, however, appears to be fragile, because splitting the population in two equal segments produces many significant differences between the two halves.

12.3.2 Reproduction at the aggregate level

The standard output file for the stability test presented in Appendix 5 also comprises the observed tourist choices of the 1998 CVO-panel. This section discusses the presence and extent of system irregularities at the aggregate level by comparing the predicted tourist trip patterns of the 1998 CVO-panel ("MERCIN-null") to the observed patterns ("98 CVO"). To this purpose, excerpts from the Tables A5.1 and A5.2 are presented, completed with calculations of the absolute and the percentage difference between the observed and the predicted tourist choices for the 1998 CVO-panel.

Table 12.1 Observed and predicted numbers of tourist trips for the 1998 CVO-panel

	98 CVO usable allocation responses *	98 CVO all members	M-null	Diff. **	%Diff ***
Av. N° of adults	1,530.00	2,836.00	2,836.00		
Av. N° of children	726.00	726.00	726.00		
For all simulation units, the average N° of (per unit):					
- Day-trips (adults only)	14.9444	****	14.6521 (.163)	****	****
- Holidays	1.9477	1.9992	2.0009 (.021)	.0017	.0850%
- Short Breaks	.6804	.7080	.7082 (.009)	.0002	.2825%
- Medium Long Holidays	.5501	.5691	.5694 (.008)	.0003	.0527%
- Extended Holidays	.3932	.4037	.4038 (.008)	.0001	.0248%
- Long Holidays	.2872	.2698	.2705 (.007)	.0007	.2595%
- Extra Long Holidays	.0368	.0486	.0489 (.003)	.0003	.6173%
- days not allocated to tourism (adults)	331.4013	-**	331.5401 (.368)	-**	-**
- days not allocated to tourism (children)	348.3072	348.3072	348.3092 (0.498)	.0020	.0006%

* For day-trips, these figures are based on the responses to the allocation task; figures on holidays, on the other hand, are based on the CVO-data.

** Diff. = M-null - 98 CVO all members

*** Difference relative to "98 CVO all members"

**** These figures cannot be presented because of the lack of information regarding the observed number of day-trips for several members of the CVO-panel.

With regard to the participation choices (Table A5.1), the comparison between the observed and the predicted tourist trip patterns of the 1998 CVO-panel is impeded by the fact that the observed participation choices are presented only for those members of the CVO-panel who's allocation responses were used to calibrate the time allocation/trip generation models (see also Table 6.2 in Chapter 6). In other words, this table only presents the observed participation choices of 1530 adults and 726 children. Table 12.1, therefore, presents an excerpt from Table A5.1 (columns "usable allocation responses" and "$\mathcal{M}$-null"), completed with the average number of holidays for the entire CVO-panel (column "all members") that allows us to compare the observed and the predicted participation choices more accurately (columns "Diff." and "%Diff."). This table shows that, overall, the average number of holidays (of different lengths) is predicted extremely accurately by the system because the largest deviation from the observed numbers amounts to only 0.6173%. On the other hand, this should not be a surprise given the adjustment factors that were introduced to the time allocation/trip generation models.

With regard to the participation and trip choices of *children*, Table 12.2 shows that even without adjustment factors, the $\mathcal{MERLIN}$-system is capable of reproducing the observed tourist choice of children with a maximum deviation of 1.80% (in case of extra long holidays).

Table 12.2 Observed and predicted participation and trip choices of 726 children (0-15 years)

	98 CVO	$\mathcal{M}$-null	Diff.	%Diff
Av. N° of children without tourist trips	114.00	113.33 (10.19)	-.67	-.59%
Av. N° of children going on holidays	612.00	612.67 (10.19)	.67	.11%
- Av. N° of holidays per unit	2.1454	2.1387 (.039)	-.0133	-.6199%
- Av. N° of SB per unit	.7500	.7457 (.020)	-.0043	-.5733%
- Av. N° of MLH per unit	.5980	.5985 (.018)	.0005	.0836%
- Av. N° of EH per unit	.4363	.4355 (.021)	-.0008	.1834%
- Av. N° of LH per unit	.3333	.3308 (.017)	.0025	.7501%
- Av. N° of ELH per unit	.0278	.0283 (.007)	.0005	1.799%
- Av. N° of days not allocated to tourism	345.1977	345.2209 (.524)	.0232	.0678%

With regard to the participation and trip choices of *adults*, it is not possible to group the members of the 1998 CVO-panel according to the participation choices based on the CVO data, because this data source does not comprise any information on day-trips. In Table A5.1, therefore, the members of the CVO-panel who's responses to the allocation model were used to estimate the time allocation

/trip generation models ("98 CVO"), are grouped based on their allocation responses. Except for data on day-trips, however, data in this column represent the 1998 CVO-data.

Table 12.3 represents the observed and predicted number of adults in each participation choice group. Again, the observed participation choices are presented only for those members of the CVO-panel whose allocation responses were used to calibrate the time allocation/trip generation models. Relative to this number of adults (98 CVO), there are 85.36% (= 100% * (2836-1530)/1530) more adult members in the entire 1998 CVO-panel. Assuming the share of each participation choice group in the "usable allocation responses" to be representative of the entire CVO-panel, "$\mathcal{M}$-null" should exceed "98 CVO usuable allocation responses" by approximately 85% for each participation choice category in Table 12.3. Indeed this applies to all groups, except perhaps for those adults making day-trips only (+89.23%).

Table 12.3 Observed and predicted number of adults (16+ years) in each participation choice group

	98 CVO usable allocation responses	$\mathcal{M}$-null	Diff.	%Diff
Av. N° of adults (16+ years):	1,530.00	2,836.00 (0.00)	1,306.00	85.36%
Av. N° of adults without tourist trips	195.00	361.05 (16.19)	166.05	85.15%
Av. N° of adults making day-trips only	133.00	251.68 (15.32)	118.68	89.23%
Av. N° of adults going on holidays only	259.00	477.70 (18.36)	218.70	84.44%
Av. N° of adults going on holidays & day-trips	943.00	1,745.57 (21.52)	802.57	85.11%

With regard to the *holiday choices* (Table A5.2), the (sizes of the) samples for the observed and the predicted tourist choices are the same and can therefore be compared straightforwardly. The predicted average length of holidays, for instance, is exactly the same as the average length of the observed holidays: 9.18 days (see Table A5.2). In the remainder of this discussion and tables, the predicted ("$\mathcal{M}$-null") and the observed ("98 CVO") tourist choices for each facet of the decision-making process are discussed in more detail for the 3562 simulation units. In this discussion, the performance of the MERLIN-system is considered "excellent" when "$\mathcal{M}$-null" deviates less than 2.5 percent from "98 CVO". The performance is "good", "quite reasonable" or "tolerable" when the differences vary between 2.5-5%, respectively 5-7.5% or 7.5-10%. More serious abnormalities occur when "$\mathcal{M}$-null" and "98 CVO" differ more than 10% (relative to "98 CVO").

Table 12.4 Observed and predicted choices of travel party

	98 CVO	$\mathcal{M}$-null	Diff.	%Diff
alone	248.00	288.26 (17.91)	40.26	16.23%
adults only	2,707.00	2,756.85 (47.18)	49.85	1.84%
Schoolchildren (6-14 years)	1,370.00	1,425.93 (36.26)	55.93	4.08%
other children (0-5 or 15+ years)	933.00	981.63 (28.88)	48.63	5.21%
party of 9+ people	717.00	750.48 (27.26)	33.48	4.67%
unknown	1,146.00	924.18 (29.66)	-221.82	-19.36%

Table 12.4 presents the observed and predicted choices of *travel party* for the members of the 1998 CVO-panel. Compared to the observed data, the predicted numbers are slightly overestimated for all choice alternatives (33-56 in absolute numbers) at the cost of the 'unknown' category that typically represents travel party choice for shorter holidays made by people who made at least 3 holidays during a particular quarter of the year (see chapter 8). Hence, even though the probabilistic decision rules greatly enhanced the ability to predict the 'unknown' parties (304%; see chapter 8), the predicted number of this travel party choice is 19.36% below the observed number. As a consequence, the predicted number of people travelling alone is also overestimated by 16.23% (relative to the observed number). For the other choice categories, the overestimation of 33-56 cases does not effect in dramatically unbalanced results (1.84% to 5.21% overestimation).

Table 12.5 Observed and predicted choices of timing

	98 CVO	$\mathcal{M}$-null	Diff.	%Diff
winter, during school holiday period	399.00	435.61 (20.73)	36.61	9.18%
winter, outside school holiday period	447.00	464.72 (24.25)	17.72	3.96%
spring, during school holiday period	944.00	1,028.88 (31.04)	84.88	8.99%
spring, outside school holiday period	1,607.00	1,563.60 (38.58)	-43.40	-2.70%
summer-begin	878.00	832.73 (23.57)	-45.27	-5.16%
summer-mid	1,082.00	1,059.38 (31.54)	-22.62	-2.09%
summer-end	652.00	637.14 (25.35)	-14.86	-2.28%
autumn, during school holiday period	322.00	334.27 (19.68)	12.27	3.81%
autumn, outside school holiday period	790.00	771.00 (27.25)	-19.00	-2.41%

In Table 12.5, the observed and the predicted choices of *season* and *school holiday periods* are presented. None of the combined timing categories is over- or underestimated by more that 10%, indicating a tolerable to excellent performance with regard to this facet of the tourist decision-making process. However, this table does show that the number of holidays during school holiday periods in the winter (9.18%) and spring (8.99%) is overestimated at the cost of the number of summer holidays (-3.17% across the three parts of this season and/or school holiday).

The performance of the *MERLIN*-system on the next facet of the tourist decision-making process, i.e. the choice of *destination*, is presented in Table 12.6. With regard to the balance between domestic and foreign holidays, the performance of the system is fairly good given the deviations of -3.11% for domestic and 3.86% for foreign holidays (the absolute differences between the observed and the predicted numbers of domestic and foreign holidays do not average out exactly because the overall number of predicted holidays slightly exceeds the observed number; see also Table 12.1). With regard to the domestic destinations, only the performance on the metaregion "Land south" is severely biased (-10.08%), especially for the western part of the metaregion (-38.15%). There are no apparent explanations for these biases.

Table 12.6 Observed and predicted choices of destination

	98 CVO	$\mathcal{M}$-null	Diff.	%Diff
* The Netherlands	3,855.00	3,735.30 (60.82)	-119.70	-3.11%
♠ metaregion "Water"	1,177.00	1,188.24 (39.98)	11.24	0.95%
- North	469.00	454.77 (22.97)	-14.23	-3.03%
- Mid	177.00	187.31 (18.93)	10.31	5.82%
- North Sea	531.00	546.16 (28.99)	15.16	2.85%
♠ metaregion "Land north"	1,491.00	1,458.83 (47.20)	-32.17	-2.16%
- Mid	580.00	549.89 (29.81)	-30.11	-5.19%
- East	485.00	527.16 (32.45)	42.16	8.69%
- North	426.00	381.76 (25.84)	-44.24	-10.38%
♠ metaregion "Land south"	976.00	877.61 (37.36)	-98.39	-10.08%
- West	465.00	287.60 (22.97)	-177.4	-38.15%
- South	511.00	590.01 (18.93)	79.01	15.46%
♠ metaregion "Other"	211.00	210.63 (20.55)	-0.37	-0.18%
* Foreign countries	3,266.00	3,392.03 (59.21)	126.03	3.86%
♣ neighbouring countries	1,656.00	1,827.89 (42.82)	171.89	10.38%
- France	656.00	706.58 (25.59)	50.58	7.71%
- Belgium & Luxembourg	423.00	474.44 (21.32)	51.44	12.16%
- Germany	392.00	445.20 (21.80)	53.2	13.57%
- United Kingdom	185.00	201.68 (15.97)	16.68	9.02%
♣ more distant countries	1,610.00	1,565.10 (41.69)	-44.90	-2.79%
- Spain & Portugal	417.00	432.98 (20.23)	15.98	3.83%
- Austria & Switzerland	322.00	273.07 (15.79)	-48.93	-15.20%
- Italy & Greece	263.00	282.76 (16.02)	19.76	7.51%
- Eastern Europe	141.00	150.38 (11.74)	9.38	6.65%
- South East Mediterranean	116.00	117.75 (10.34)	1.75	1.51%
- Scandinavia	93.00	99.11 (10.15)	6.11	6.57%
- Other	258.00	208.08 (13.64)	-49.92	-19.35%

With regard to foreign destinations, the number of holidays to neighbouring countries is overestimated by 10.38%, whereas the number of holidays to more distant destinations is reproduced quite well (-2.79%). In other words, the slightly overestimated number of foreign holidays mainly affects the reproduction of holidays to neighbouring countries, in particular Belgium and Luxembourg (12.16%) and Germany (13.57%). Within the more distant destinations, only the predicted number of holidays to Austria and Switzerland (-15.20%) and "other" destinations (-19.35%) are severely underestimated compared to the observed numbers. Again, there are no apparent reasons for these abnormalities.

Table 12.7 Observed and predicted choices of accommodation

	98 CVO	*M*-null	Diff.	%Diff
hotel	2,347.00	2,440.68 (53.30)	93.68	3.99%
non-permanent - owned by traveller	998.00	921.24 (30.87)	-76.76	-7.69%
permanent - owned by traveller	1,170.00	975.52 (37.15)	-194.48	-16.62%
other	2,606.00	2,789.89 (54.92)	183.89	7.06%

Table 12.7 presents the observed and predicted choices of *accommodation*. This table shows that the *MERLIN*-system has some difficulties in reproducing the number of holidays that are spent in privately owned accommodations, in particular accommodations with a permanent location (-16.62%). These abnormalities can be explained - at least in part - by the underrepresentation of 'unknown' travel parties that was discussed earlier because this condition is frequently involved in the decision rules that favour privately owned tourist accommodations (see Table 11.1, R48, R49, R75 and especially R85-88).

Table 12.8 Observed and predicted choices of transport mode

	98 CVO	*M*-null	Diff.	%Diff
car	5,335.00	5,321.90 (63.27)	-13.1	-0.25%
aeroplane	868.00	812.95 (29.47)	-55.05	-6.34%
alternative, of which	918.00	992.48 (32.35)	74.48	8.11%
- bus	416.00	449.74 (21.04)	33.74	8.11%
- train	260.00	274.26 (16.16)	14.26	5.48%
- other	242.00	268.48 (16.45)	26.48	10.94%

In the case of *transport mode choices* (see Table 12.8), the *MERLIN*-system is able to reproduce the choices between car, aeroplane and alternative modes quite reasonably. With regard to the alternative modes, only the other modes are overestimated (slightly) more than 10%.

The second last facet of the tourist decision-making process is the *date of departure*. Table 12.9 shows that the *MERLIN*-system's ability to reproduce this facet at the aggregate level is good.

Table 12.9 Observed and predicted choices of the date of departure

	98 CVO	$\mathcal{M}$-null	Diff.	%Diff
Friday/Saturday	4,054.00	3,922.80 (61.61)	-131.2	-3.24%
other day of the week	3,067.00	3,204.53 (52.75)	137.53	4.48%

Finally, Table 12.10 presents the observed and the predicted choices of *expenditures* for both domestic and foreign holidays. The overall predicted average expenditures per holiday are good. Broken down according to the trip destination, the performance for domestic holidays it is tolorable, whereas for foreign holidays it is excellent. Broken down further according to duration, however, the expenditures on short breaks (and also foreign medium long holidays) are severely overestimated, whereas for extra long holidays, the expenditures are seriously underestimated. These biases can, at least in part, be attributed to the approach that was used to derive decision-making rules for the choice of expenditures[39].

Table 12.10 Observed and predicted choices of expenditures for domestic and foreign holidays broken down according to the duration of the holiday

	98 CVO	$\mathcal{M}$-null	Diff.	%Diff
* holiday	713.03	744.50 (11.51)	31.47	4.41%
♠ domestic holiday	268.42	289.04 (4.75)	20.62	7.68%
- domestic SB	155.87	180.01 (3.87)	24.14	15.49%
- domestic MLH	299.24	321.02 (8.35)	21.78	7.28%
- domestic EH	468.70	491.47 (21.47)	22.77	4.86%
- domestic LH	617.52	559.49 (33.17)	-58.03	-9.40%
- domestic ELH	803.92	541.43 (75.17)	-262.49	-32.65%
♣ holiday abroad	1,237.82	1,246.08 (20.40)	8.26	0.67%
- SB abroad	386.51	507.23 (25.39)	120.72	31.23%
- MLH abroad	846.23	1,038.07 (35.43)	191.84	22.67%
- EH abroad	1,393.99	1,503.14 (40.25)	109.15	7.83%
- LH abroad	1,761.05	1,685.43 (45.97)	-75.62	-4.29%
- ELH abroad	3,242.39	2,043.39 (106.61)	-1199	-36.98%

More specifically, with regard to the *extra long holidays*, chapter 11 already discussed that the frequency of these extra long holidays proved to be too low to distinguish them from long holidays (in the decision rules). In addition, in the simulation process, the upper limit of the 10^{th} segment of the expenditure alternatives was set at NLG 4750 to ensure that the predicted mean matches the

[39] Several alternative approaches were tested, including a Poisson regression model and a decision rules that indicated average expenditures directly (see also chapter 11), but these approaches either increased the discussed biases, introduced other (more severe) biases and/or violated model assumptions.

observed mean. As a consequence, several extra long holidays may not be classified in the highest segment, and, in addition, those that are classified in this segment, can never be more expensive than NLG 4750. In combination, these two effects result in relatively low predicted expenditures on extra long holidays.

In contrast, *short breaks* are relatively cheap and, based on the decision rules, should be classified in the first segment more frequently. Based on the observed and the linear means of the first segment (NLG 43.15 respectively NLG 27.50) there was no reason to lower the predicted mean of this category. In fact, the (linear) mean should have been increased to meet the observed mean for this segment. In any case, as a result of the assumed linear distribution of expenditures in this segment, short breaks, and especially the relatively frequent trips to privately owned accommodations with a permanent location, that are often on the lower end of this segment, which may explain the relatively high predicted expenditures.

12.3.3 Reproduction at the individual level

In addition to the comparison between the predicted and the observed data at the aggregate level, the simulation outcomes of *MERLIN*-null with regard to the number of holidays and day-trips, the destination choices, the duration choices and expenditures were also compared to the observed choices at the individual level. In this test, the predicted values for the indicators of *MERLIN*-null (i.e. the average value across 200 runs) were regressed onto the observed values for these indicators (i.e. the 1998 CVO-values). If the slope for the predicted value approximates one (i.e., a change in the observed value is followed by an equal change in the average predicted value), and the intercept does not differ from zero significantly (i.e., there is no systematic over- or underprediction), the *MERLIN*-null is able to capture the tourist choice process at the individual level adequately. Table 12.11 presents the results of the regression analyses.

For the annual number of day-trips, Table 12.11 presents the relationship between the observed and the predicted number of day-trips for the 1530 members of the 1998 CVO-panel whose responses to the allocation task were used in the model estimations. The results show that the intercept does not differ from zero significantly (at the 5% level). The slope of the regression line, on the other hand, is not steep enough (o.813), but it does not differ significantly from unity at the 1% level.

With regard to the average number of holidays, both the overall number of holidays and the number of domestic holidays are reproduced very well at the individual level. The slope of the regression line for the average number of foreign holidays, on the other hand, differs from one more seriously, while the intercept differs from zero at the 5% level. The value of the intercept, however, is very small (0.07007), and even the value of the slope (0.889) does not cause any serious problems.

Table 12.11 Regression analysis for the observed and predicted values of several indicators at the individual level (Observed value = *intercept* + *slope* * Predicted value)

	intercept (*t*-value*) + *slope* (*t*-value**)	R / R^2
# day-trips (N=1530)	2.364 (1.67) + .813 (-2.22)	.240 / .058
# holidays	.112 (1.66) +.943 (-2.01)	.487 / .237
# domestic holidays	.005358 (.111) + 1.027 (.803)	.455 / .207
# holidays abroad	.07007 (2.04) + .889 (-3.53)	.428 / .183
# SB	.06136 (1.62) + .913 (-2.44)	.394 / .156
# MLH	.113 (3.15) + .802 (-3.50)	.321 / .053
# EH	-.008067 (-.256) + 1.020 (.271)	.225 / .051
# LH	.03136 (2.25) + .881 (-2.74)	.322 / .104
# ELH	.02602 (6.31) + .461 (-18.4)	.255 / .065
Av. duration of holidays (in days)	1.365 (3.63) + 1.009 (.171)	.306 / .094
# days allocated to holidays	2.362 (4.55) + .870 (-5.31)	.512 / .262
Annual expenditures on holidays	-129.25 (-2.23) + 1.044 (1.29)	.456 / .206

* *t*-value with respect to deviation from 0; i.e., (estimated slope)/(standard deviation);

** *t*-value with respect to deviation from 1 (unity); i.e., (estimated slope - 1)/(standard deviation); a negative *t*-value indicates a slope value smaller than unity and a positive *t*-value indicates a slope value larger than unity

On the subject of duration choices, Table 12.11 shows that the MERCIN-system is able to capture the numbers of short breaks and extended holidays very well at the individual level. For medium long, long and extra holidays, the intercepts are slightly too high (0.113, 0.03136 and 0.02602), but do not cause any serious problems. The slopes of these regression lines, on the other hand, are not steep enough (0.802, 0.881 and 0.461). In particular in the case of extra long holidays, the slope (0.461) indicates a serious overrepresentation of the number of holidays for the majority of units. This distortion should probably be attributed to the low frequencies of extra long holidays and the relatively high number of tourists that do not participate in extra long holidays. These data particularities seriously violate the assumptions of regression analysis. Table 12.12, therefore shows a comparison of the mean number of extra long holidays according to the observed frequency. This table shows that, on average, MERCIN predicts higher trip frequencies for the members of the 1998 CVO-panel that made one or more extra long holidays (compared to those who did not make any extra long holidays).

With regard to the average duration of holidays, the annual number of days allocated to holidays and the annual expenditures on holiday, Table 12.11 indicates the slopes of the regression analyses to approach one reasonably well (although the slope for the annual number of days allocated to holidays tends to be slightly too low (0.870)). The intercepts for these indicators, however, differ significantly from zero: 1.365 for the average duration of holidays, 2.362 for the average number days of allocated to holidays annually and -129.25 for the annual expenditures. Since the slope for the average number days of allocated to holidays is smaller than one (0.870), the structural overestimation indicated by the intercept is compensated. This explains why the predicted average length of holidays at the aggregate level is (exactly) the same as the average length of observed holidays: 9.18 days (see Table A5.2). The intercept for the annual expenditures on holidays is significantly (at the 5% significance level) smaller than zero. The explanation for the negative value of this intercept can be found in the number of people who did not make any holidays according 1998 CVO-data. In the simulation process, the majority of these simulation units will, at some point in the 200 runs, have been "assigned" one or more holidays because the process uses probabilities and random numbers. The average expenditures on holidays for these simulation units across 200 runs, therefore, will always be larger than zero. Accordingly, the regression line will cut the horizontal predicted-axis at a positive value, and the vertical observed-axis at a negative value. This negative value of the intercept is, to some extent, compensated by the value of the slope (1.044); the slope, however, does not differ from unity significantly.

Table 12.12 Average predicted number of extra long holidays (ELH) according to the observed number of ELH

Observed # of ELH	Predicted # of ELH	Number of observations
0	.04327	3405
1	.1436	144
2	.2670	10
3	1.2217	3
Total	.04895	3562

Table 12.11 also presents the adjusted value of the slope (beta or R), and the part of the variance that is explained by the model (R^2, i.e. the dispersion of the observations around the (linear) regression line). This part of the table indicates that, although the relationship between the observed and the predicted values at the individual level is often quite reasonable, the part of the variance for each indicator that is explained by the *MERLIN*-system is relatively low: the explained part of the variance exceeds 20% for 4 indicators only.

12.3.4 Conclusion and discussion

This section discussed several validation issues of the MERCIN-system. More specifically, this section discussed the stability of the simulation outcomes and the reproduction at the aggregate and the individual level for the most important baseline simulation, i.e. the default reference project "MERCIN-null" that represents the simulated tourist trip patterns of the 3562 members of 1998 CVO-panel without any form of upgrading and/or ageing. With regard to the stability of the simulation outcomes it was concluded that the results were satisfactorily both in terms of the convergence of the mean value across consecutive simulation runs and the reproductive power for identical input conditions. The homogeneity of the population, however, appeared to be fragile, because splitting the population in two equal segments produces many significant differences between the two halves.

At the aggregate level, MERCIN is able to reproduce the original data quite reasonably. For the participation choices this was expected given the weighting schemes that were introduced to this part of the system. With regard to the facets of the tourist decision-making process for holidays, the majority of indicators proved to be satisfactory. However, several system abnormalities appeared. First, the number of "unknown" travel parties was underestimated by 19.36%, which also caused a serious underrepresentation of the number of holidays in tourist accommodations with a permanent location that are owned by the traveller (-16.62%). Second, there were some biases in the prediction of several destinations. More specifically, visits to the Dutch metaregion "Land south", Austria and Switzerland and "other more distant countries" were underestimated, while visits to neighbouring countries such as Belgium and Luxembourg and Germany were overestimated.

Third, the predicted number of holidays by "other" alternative transport modes exceeded the observed number slightly more than 10%. Also, the expenditures on short breaks were severely overestimated (15-31%), while those for extra long holidays were seriously underestimated (-33 to 37%). In contrast to the first three abnormalities, the latter biases can, at least in part, be attributed to a mismatch between the data characteristics and the selected modelling approach. However, none of the alternative modelling approaches that were tested proved to describe the data more accurately. Regardless of the source of these system irregularities, however, the interpretation of impact analyses should take into account these biases.

Finally, the MERCIN-system was assessed in terms of its ability to reproduce tourist decisions at the individual level. Regression analyses was used to test the relationship between predicted and the observed data at the individual level with regard to the number of holidays and day-trips, the destination choices, the duration choices and the expenditures. These relationships proved to be quite reasonable for most indicators. However, for the number of extra long holidays, the slope of the regression line proved to be very small due to data-particularities. In addition, the intercepts for the average duration of holidays, the total number of days allocated to holidays and the annual expenditures differed significantly from zero. Finally, although the relationship between the observed and the predicted values at the individual level was adequate, the part of the variance for each indicator that is explained by *MERLIN* proved to be rather low.

Overall, the validity of *MERLIN* proves to be quite reasonably, although the interpretation of impact analyses should take into account several biases. Before demonstrating three of these analyses, the next section will first discuss the options that are available for impact analysis in *MERLIN*.

12.4 The policy options available in the *MERLIN*-system

Now that the *MERLIN*-system has been constructed and validated, it can be put to use. Appendix 4 presents the user manual. Step-by-step, this manual guides the user through the various stages of defining a project, running a project to assess the impact of a scenario for the future on tourist trip patterns, redefining a project, comparing it to other projects, and so on. Basically, defining a project consists of three important steps. First, each project has to be named and described in order to distinguish it from other projects. Secondly, the project settings have to be determined. This includes the number of simulation runs, the reference project and the decision whether or not data-bases should be produced during the simulation process. Last but not least, the scenario for the future has to be determined. Based on the latter settings, a population of units is composed whose tourist choices are simulated to assess the impact of the defined scenario for the future. Before discussing the results of three of these projects (sections 12.5 through 12.7), this section first details the available “buttons” for changing the characteristics of the population. Also, this section discusses how project settings are implemented in the simulation process to represent the defined scenarios for the future (i.e., how the characteristics of the 1998 CVO-sample are modified to produce the population of simulation units).

Given the conceptual and empirical building blocks, there are three types of policy options available in the *MERLIN*-system, including:

(1) changes at the level of tourist trip programs that were obtained using stated preference techniques (see chapters 6 and 7);
(2) changes in the general preferences for the available transport model alternatives (see chapter 10); and
(3) changes in the characteristics of the population (to simulate, for instance, the impact of the ageing of the population, increased income-levels, or the increase of people owning a tourist accommodation).

The following three subsections discuss these options in detail, including the way these changes are implemented in the system.

12.4.1 Reallocation scenarios

Given the adopted data-collection approach (see chapter 6), data were available on the number of overnight tourist trips for all 3562 members of the 1998 CVO-panel. In addition, the data also included the number of day-trips and some information on re-allocation behaviour as a result of relieving or reinforcing constraints in the tourist's decision-making context of those members aged 16 years and older. In total, 13 "group-specific-scenarios" were tested on this "adult sample" (see Figure 6.3), 7 of which were included in the time allocation/trip generation model (see chapter 7). The re-allocation scenarios that are available in the MERCIN-system are presented in Figure 12.1. If one or more of these scenario's is (are) selected in a project, the population of simulation units will be restricted to the relevant socio-economic group(s). The percentage of units from the particular socio-economic group(s) that will be confronted with these scenarios can be varied (except for scenario I/group 3, because in this scenario, the new conditions of the Student Public Transport Ticket applies to all students). The actual assignment of the scenarios to the members of the population of simulation units is determined using random numbers. If, for instance, the desired percentage of people with 12 extra free days is 9.5%, a simulation unit working at least 30 hours per week will be assigned the 12 extra free days when the random number generator returns a value between 0 and 0.095. This assignment procedure is repeated for each simulation run to avoid biases in the results due to random effects.

If a simulation unit is assigned a particular scenario, the parameters for this scenario in the time allocation/trip generation models will be activated. In addition, in some case the condition states of these simulation units will be adjusted to also represent these changes in tourist choices that are represented by rulebases.

Group 1: People working at least 30 hours per week (in 1998)
Scenario A: Due to new collective labour agreements you had had 12 extra free days; your salary had been the same;
Scenario B: Due to new collective labour agreements you had had 12 free days less; your salary, however, had increased by 5%;
Scenario E: Due to new collective labour agreements you had worked 4 days a week; Friday had been your fixed free day; your salary had been the same.

Group 2: People working less than 30 h.p.w. with a partner working at least 30 h.p.w.
Scenario F: Due to new collective labour agreements your partner had had 12 extra free days; his/her salary had been the same;
Scenario G: Due to new collective labour agreements your partner had had 12 free days less; his/her salary, however, had been increased by 5 %.

Group 3: Students and pupils
Scenario I: In 1998 you had been allowed to use your Student Public Transport Ticket as it was introduced in 1991, viz. free use of all public transport modes within the Netherlands on all days.

Group 4: Pensioners, early retirement (Dutch: VUT) and people of independent means
Scenario J: Due to changes in the Dutch tax and social legislation, your net household income had been 15% higher in 1998.

Figure 12.3 The re-allocation scenarios available in the *MERLIN*-system.

More specifically, the following adjustments are carried through for the scenario's:

- Scenario A (12 extra free days): the variable "Freed" (the number of free days in 1998) is raised by 12, with a maximum value of 98;
- Scenario B (12 free days less; income raised by 5%): the variable "Freed" (the number of free days in 1998) is reduced by 12 days, with a minimum value of 20 (the legal minimum); in addition, for 20% of the simulation units to whom this scenario is assigned, the household income level ("Inc", levels: 1-8) is raised one level, with a maximum value of 8;
- Scenario E (4-day working week, Fridays off): the value of the variable "Wrkhr" is set at 32 working hours per week;
- Scenario F (12 extra free days for the partner): no adjustments;
- Scenario G (12 free days less for the partner; partner's income raised by 5%): for 20% of the simulation units to whom this scenario is assigned, the household income level ("Inc", levels: 1-8) is raised one level, with a maximum value of 8;
- Scenario I (improved Student Transport Ticket): no adjustments;
- Scenario J (15% more income for seniors): for all simulation units to whom this scenario is assigned, the household income level ("Inc", levels: 1-8) is raised one level, with a maximum value of 8.

Note that these condition states are adjusted *only after* the trip generation stage is executed, because the scenarios are represented in the trip models by parameters.

12.4.2 Aggregate (nation-wide) preferences for modes of transport

The second option to specify scenario's for the future is related to aggregate transport mode preferences. In section 5.2 it was argued that, even though transport mode choices in the MERCIN-system are taken conditional upon the destination choices, the latter choices will be deliberated while taking into account the condition of the transport system. Based on these insights, decision rules for destination choices (see Chapter 10) were modelled to "look ahead" in order that changes in the transport system may affect destination choices. This objective was attained by including condition variables that express (on a scale from 0 to 100) the tourist's general propensities to select the most important transport mode alternatives for holidays, i.e. "car", "aeroplane" or "alternative". Each variable expresses the *individual's general propensity* to select one of these three mode alternatives, and together the three propensity variables add up to 100 for each individual. These propensities are related to the individual rather than the holidays because they represent the share of each transport mode over all holidays of an individual. In section 10.2 a Multinomial Logit (MNL-) model was calibrated to describe for each individual the general propensity to select a particular transport mode dependent based on (1) the aggregate marginal distribution of the transport mode alternatives over all travellers and over all observed holidays; (2) the individual's annual number of holidays; and (3) several personal and household characteristics. By changing the aggregate marginal distribution over the modes (e.g. due to increasing costs of fuel), the user can specify projects that assess the effects of these changes on tourist trip patterns. These changes, then, are carried through by using the user-specified aggregate marginal distribution of the transport mode alternatives in the algorithm of the calibrated MNL-model[40].

12.4.3 Composition of the population

Third and finally, the MERCIN-system offers the opportunity to specify changes in the composition of the population. These changes are implemented by *statically ageing* and/or *uprating* of the population because the system is based on

[40] If the user does not specify an aggregate marginal distribution for the transportation modes, the default values that were used to calibrate the MNL-model in section 10.2 are used, i.e., 74.9 for "car", 12.2 for "aeroplane" and 12.9 for "alternative".

information on simulation units at a certain point in time (see also chapter 4). Experiments are therefore conducted by systematically changing the input-characteristics of the system. In order to assess the impact of the greying of the population, for instance, the relation between the different age groups can be changed. Or, the impact of a promotion campaign and/or a price reduction for tourist accommodations with a permanent location can be assessed by increasing the number of people with these accommodations (assuming this would be the effect of the promotion and/or price reduction).

Basically, there are two ways to implement these population changes (see also chapter 4). First, a cross-section can be *uprated* by changing the characteristics (conditions) of the simulation units. Secondly, the share of units that meet the user-defined conditions can be adjusted to *statically age* the sample. Uprating (an appropriate subset of) simulation units is possible only when the characteristic under consideration is not (or hardly) related to other sample conditions. These characteristics include, for instance, the possession of cars, skis or tourist accommodations[41]. In this case, if the number of units with certain characteristics is to be increased, units that do not have this characteristic have a certain probability of obtaining it. Oppositely, when the number is to be decreased, units that possess a certain characteristic have a certain probability of losing it.

If, on the other hand, characteristics are (heavily) correlated with other population characteristics, it is more appropriate to change the share of units that meet the user-defined characteristics to control for these correlations. Changing the working situation, for instance, will also affect the household income. Similarly, the "greying" the population by increasing the proportion of elderly people will also affect, for instance, the household composition ('empty nests'). In this case, conditional probabilities should be defined that specify the probability for a member of the original sample to become a member of the population that will be simulated. An example may clarify this principle. Let the sample of respondents on which the system is based consist of 1000 individuals, 47.5% of which have at least one child in the household and 52.5% of which are living in "childless" households. If the user is interested in the effects of an increase in the number of "childless" households to 55%, respondents that belong to the latter segment have a probability of 104.76% (550/525 * 100%) of becoming part of the population that will be simulated. In other words, they have a probability of 4.76% of being

[41] This way of implementing changes is also applied to the condition "tied to certain holiday periods due to working obligations" because it is plausible that this condition is tied to the type of job (and not, for instance, the income or educational level). Also, people without a paid job can be tied to certain periods due to, for instance, voluntary work.

represented in the simulation population twice. Conversely, respondents in households with children have a chance of only 94.73% (450/475 x 100%) of being represented in the simulation population. Once the conditional probabilities have been calculated, the simulation units are "selected" using random numbers.

The above principle can be extended to generate conditional probabilities based on two or more changing characteristics. This applies to cases in which the user specifies two or more changes, but also when a characteristic is strongly related to another characteristic (that should not change). However, the present version of MERCIN-system reduces the complexity of these calculations by assuming independence and implementing these conditional probabilities sequentially[42] (the order of implementation is indicated by the Arabic numerals in the policy wizard). In order to reduce the effects of this assumption, one exception is made with regard to the relationship between work and gender. Table 12.13 shows that in the Netherlands, men work more frequently than woman. More importantly, men often work full-time while women work part-time. This means that if the number of part-time (< 30 h/w) and/or non-working people would be increased by changing the appropriate share(s) of simulation units, women would be overrepresented seriously. To forestall this problem, MERCIN offers the opportunity to specify gender-specific changes to the work situation.

Table 12.13 The relationship between gender and the work situation of adults (≥ 16 years) based on the 1998 CVO-data (N = 2836)

	Work situation			Total
	Paid job ≥ 30 h/w	Paid job < 30 h/w	No paid job	
Female	362 (23.7%)	341 (22.3%)	825 (54.0%)	1528 (100%)
Male	816 (62.4%)	61 (4.7%)	431 (33.0%)	1308 (100%)
Total	1178 (41.5%)	402 (14.2%)	1256 (44.3%)	2836 (100%)

Finally, it should be noted that, similar to the re-allocation scenarios, the assignment procedure for uprating or statically ageing the population is repeated for each run to avoid biases due to random effects. The *average* composition of the population is presented in the standard output-file for each project.

12.5 Demonstration projects

The previous sections have discussed the validity and the possibilities of MERCIN in terms of impact analysis for three types of scenarios for the future.

[42] This assumption is partially justified by the observation that, in practice, the majority of projects will change only one, or at the most two characteristics of the population.

For each of these types, this final section demonstrates one example. Demonstration A shows an example of the first reallocation scenario where 10% of the full-time (> 30 hours per week) working members of the population is granted 12 extra free days. Secondly, section 12.5.2 discusses the effects of changes in the aggregate preferences for various modes of transport. Demonstration C, finally, discusses the effects of the ageing of the population on tourist choices.

The simulation results of these demonstration projects are presented in Appendix 6, Tables A6.1 through to A6.7 inclusive, where the simulation results of the demonstration projects are compared statistically to their respective baseline simulations. The number of prediction runs is set to N = 25 for both the demonstration and the baseline simulations. In section 12.3.1 it was shown that two identical simulation projects will still produce 3-5 statistically significant differences on the indicators. In addition to the “standard” comparison between the demonstration project and the baseline simulations, therefore, a test for stability of the simulation results was conducted to exclude “significant” changes between the baseline and the demonstration project that are based on chance. In this test, the demonstration project is simulated again using the same project and simulations settings. The simulation results of this second (identical) project are also compared to the *baseline simulation*. In Appendix 6, the results of this second test (the column labelled ‘S2’) are presented in the column on the right-hand side of the first test for changes between the demonstration project and the baseline scenario (the column labelled ‘S1’). Differences between ‘S1’ and ‘S2’ indicate the alleged significant change to be based on chance; if the differences between the baseline simulation and the scenario for the future are significant for both projects, the cells are framed to mark this correspondence (in all cases, the sign of the change(s) proved to be consistent).

12.5.1 Demonstration A: 10% of the full-time working members of the population are granted 12 extra free days per year

The first demonstration project is an example of the re-allocation scenarios that were discussed in section 12.4.1. The project assesses the impact of an increase of 12 days in the annual number of free days for 10% of the people working at least 30 hours per week (scenario A; see Figure 12.3). The impact of these changes are assessed *for this segment of the population only*. The segment under consideration comprises 1192 members, i.e. 33.5% of the members of the 1998 CVO-panel.

First, the appropriate baseline simulation is created by simulating (N = 25) the tourist choices for the 1192 adults in this segment of the population for the “no change"-situation. Next, on average 10% of the segment is granted 12 extra free days per year in each simulation run, and the results of these simulations (N = 25) are compared to the baseline simulation. Table A6.1 presents the resulting participation choices for this scenario, while Table A6.2 presents the resulting holiday choices.

OBSERVED IMPACT AND INTERPRETATION OF THESE CHANGES

The proposed increase in the annual number of free days results in an increase in the number of people that only go on holidays of almost 6% for this segment of the population. This observation corresponds with the parameters in the participation choice model for adults. The decrease in the number of people pursuing both holidays and day-trips is considerable but not reliable, and should be attributed to chance.

With regard to trip frequencies, the proposed increase in the annual number of free days results in an increase in the average number of holidays per person in this segment (+0.56 holidays p.p.p.y, i.e. +2.68% relative to the baseline; in absolute numbers, approximately 67 extra holidays are made by the members of this segment). In particular the number of extra long holidays (+10.18%), and to a lesser extent also short breaks (+4.21%) and medium long holidays (+3.15%) become more popular due to the increased number of free days. There is no effect on the number of day-trips. Apparently, the working adults are satisfied with regard to their day-trip frequencies, and allocate (at least a part of their) extra free time to holidays. The number of days *not* allocated to tourism, however, decreases only slightly (approximately -0.3 days per year), and this decrease is only significant in the demonstration project ('S1') and not in the test project ('S2'). This accords with the (non-significant) decrease in the average duration of holidays (see Table A6.2).

In terms of the holiday choices, Table A6.2 shows that the number of holidays without travelling companions and/or with large groups will not increase. In contrast, holidays with adults only (+2.03%), with very young (≤ 5 years) and/or older (≥ 15 years) children (+5.17%) and with unknown travel parties (+5.96%) will become more popular due to the proposed changes. The latter impact indicates that people who already pursue holidays frequently and/or possess a tourist accommodation with a permanent location, are also inclined to make more holidays when they are granted more free days because these conditions increase the probability of unknown travel parties (see chapter 8).

With regard to timing choices, the extra free days for 10% of the segment will result in a relatively strong increase in the number of spring holidays, both during (+4.39%) and outside (+2.80%) school holidays periods. There are also important increases in the number of holidays during the other seasons (especially autumn and the first and last part of the summer), but these changes are not sigificant. In other words, the increase in the number of free days will (continue to) add to the importance of the early (and late) season for this segment of the population.

The analyses also show that the increase in the number of free days will advance domestic holidays more than foreign holidays both relatively (+3.15% vs. +2.28%) and absolutely (+35.84 vs. +31.32 holidays per year). This is an important conclusion for the Dutch tourism industry, because the segment under consideration is traditionally more focussed on foreign holidays than other segments of the population. Within the Netherlands, the metaregion “Land north” will take particular advantage of the increase in the number of free days. Outside the Netherlands, neighbouring countries, in particular the United Kingdom, will profit from the proposed changes. The impact on destination changes is in accordance with the increase in relatively short holidays (short breaks and medium long holidays). Apparently, the increase in extra long holidays, that are more likely to be spent in more distant areas, is too weak to result in (significant) changes in these more distant destinations.

In terms of accommodation choices, all accommodation types profit from the increased number of free days (and the resulting extra holidays), but only hotels (+2.48%) and “other” accommodations (+2.43%) do so significantly. This also applies to the available means of transport, where only the use of cars (+2.70%) shows a significant increase for this segment of the population.

Members of the segment under consideration do not leave on Fridays and Saturdays more than average (55.45% (see Table A6.2) against a national average of 56.93% (based on the 1998 CVO-data)), and the proposed increase in the number of free days does not change this relation. Apparently the extra free days are not used to pursue holidays during different parts of the week. Perhaps these observations can be explained by timing constraints of travelling companions and/or members of the household.

Finally, the average expenditures on holidays for the segment under consideration do not change significantly due to the proposed changes. The total expenditures on holidays by this segment are therefore increased only by the increase in the number of holidays.

Based on the above discussion, it can be concluded that the increase in the annual number of free days for 10% of the full-time working adults will partially be allocated to holidays while the number of day-trips remains unchanged. The analyses also show that the position of the early season will be strengthened. Also, nearby destinations including both the Netherlands and neighbouring countries profit more than average from these changes, probably due to an increase in the number of shorter holidays (< 9 days). The increase in the number of extra long holidays (29 days and longer), on the other hand, does not result in (significantly) increased numbers of more distant holidays. Other remarkable changes include the increases in holidays by adult only, parties with “other” children and “unknown” travel parties, hotel-based holidays, holidays in/to “other” accommodations, and the increased use of cars. Changes in the expenditures on tourism, finally, will only result from the increased number of holidays because the average expenditures on holidays remain unchanging for this segment of the population.

12.5.2 Demonstration B: Changes in the preferences for transport modes at the aggregate level

The second type of policy options allows the user to change the aggregate preferences for the use of cars, aeroplanes and alternative means of transport. These changes are carried through by calculating the individual propensities to use these transport modes using a Multinomial Logit (MNL-) model (see chapter 10). These individual propensities can be part of the decision rules for selecting destinations and transport modes. In the previous chapters, it has already been hinted for several stages in this process that the carry-over of proposed changes in the aggregate preferences for transport modes did not work out as expected. More specifically:

(1) The effect of the aggregate marginal distribution of the transport modes on the individual propensities to select these modes (as calculated using the MNL-model) is modest because, in the MNL-model, the parameter for this variable amounts to 0.0118 only (see chapter 10);

(2) The effect of the individual propensities on the destination choices is also modest because the condition variables representing these propensities do not occur frequently in the decision rules for destinations choices; moreover, when the condition variables under consideration occur in the decision rules for holiday destinations, they are never among the most important conditioning variables (see chapter 10); and finally

(3) The individual propensities to select particular transport modes do not effect the actual mode choices at all, because the condition variables representing these propensities do not occur in the decision rules for transport modes;

the absence of these conditions in the decision rules for transport modes implies that user-defined changes in the aggregate preferences for these transport modes will only affect the mode choices through the effect on destination choices, and since chapter 10 has already indicated that shifts in transport mode propensities are not likely to have significant effects on destination choices, the carry-over effect on transport mode choices is probably nil.

To test the impact of changes in the general propensities to select particular means of transport, five of the most extreme scenarios were formulated within the allowed ranges for each transport mode. More specifically, the aggregate preferences for car-aeroplane-alternative were set at 70-25-5, 70-0-30, 95-0-5, 50-20-30 and 50-30-20. As expected, none of these extreme changes caused any significant changes on the destination and/or transport mode choices. By way of illustration, the results of the 70-0-30-scenario are presented in Appendix 6 (Tables A6.3 and A6.4). Although not significantly, in Table A6.4, for instance, the preferences for aeroplanes even increases while the scenario is set to exclude this mode altogether.

Based on these tests, it can be concluded that the decision rules for destination choices do not “look ahead” as expected. The technical reason in terms of the models’ structures is evident. The “reluctance” of the destination and transport mode choice models to integrate the individual transport mode preferences in their condition structures, on the other hand, is less obvious.

12.5.3 Demonstration C: Ageing of the population

The third and final demonstration project is an example of the scenarios that change the composition of the population (see section 12.4.3). It was decided to change the composition of the population according to age, because the greying of the population presents one of the most clear and challenging demographic changes that await the European community in the near and medium-term future.

Table 12.14 presents the expected development of the Dutch population according to age between 2000 and 2050. During this period, the population size will increase from 15.9 million people in 2000 to 18.0 million in 2050. At the same time, the share of retired people will grow from 13.57% to 21.64%, while the share of adults younger than 65 years will decrease from 62.02% to 55.64%; the share of young people will only slightly decrease from 24.41% to 22.72% in 2050.

Table 12.14 Expected development of the Dutch population according to age 2000-2050 (inhabitants $\times$ 1000)

Year	0-19 years	20-64 years	65+ years	Total
2000	3,873.0 (24.41%)	9,838.5 (62.02%)	2,152.4 (13.57%)	15,864.0
2010	4,079.8 (24.19%)	10,286.2 (60.99%)	2,498.6 (14.82%)	16,864.5
2020	4,010.3 (22.93%)	10,262.1 (58.67%)	3,219.7 (18.41%)	17,492.1
2030	4,008.2 (22.37%)	10,062.7 (56.16%)	3,847.9 (21.47%)	17,918.9
2040	4,113.5 (22.78%)	9,806.5 (54.31%)	4,138.0 (22.91%)	18,058.0
2050	4,090.3 (22.72%)	10,018.3 (55.64%)	3,896.3 (21.64%)	18,004.9

Based on CBS (2001)

In this demonstration project, the year 2020 is selected as the planning horizon. Next, only the *shares* of the various age groups within the population are adjusted because in the present version of the MERCIN-system it is not possible to increase the *size* of the population. The assumed changes for the age-groups in MERCIN for the year 2020 are presented in Table 12.15. The share of the oldest group in MERCIN (56+ years) is assumed to follow the development of the oldest group in Table 12.14 (an increase of almost 36% relative to the share in the year 2000). For the two youngest age-groups, only very small decreases in their shares are assumed, while the decrease is the strongest for the age group between 31 and 55 years. The simulation results (N = 25) are compared to the standard reference project MERCIN-null (N = 200) and presented in Table A6.5 (composition of the population), Table A6.6 (participation choices) and Table A6.7 (holidays choices).

Table 12.15 Proposed development of the Dutch population according to age 1998-2020 (shares of the age-groups)

Age group	1998 (CVO-data)	Proposed 2020	Simulation results
0-15 years	20.4 %	19.5 %	19.51%
16-30 years	15.6 %	14.0 %	13.92%
31-55 years	42.1 %	36.5 %	36.61%
56+ years	22.0 %	30.0%	29.95%

AVERAGE COMPOSITION OF THE POPULATION

As mentioned before, the assignment procedure for statically ageing the population is repeated for each simulation run to avoid biases in the results due to random effects. The *average* composition of the aged population (including the standard deviation) is presented in Table A6.5 in Appendix 6. This table shows that the simulated population approximates the proposed shares of the various age-segments adequately (these figures are also presented in the column on the right in Table 12.15).

In addition to these changes, the composition of the population is also changed with regard to other socio-economic and –demographic variables because age is related to these variables. First, the income composition is shifted towards the lower income categories. This is in accordance with several scenarios studies in the Netherlands, for instance in the field of transport (MuConsult, 1997). This shift is, at least partially, caused by a decrease in the number of people with a paid job (from 1580 in 1998 to 1431 working members of the population in the year 2020).

Next, there are some minor changes in the education levels and the province and school holiday region of residence, while the composition with regard to gender remains (almost) unchanged. With regard to civil state, the numbers of married, divorced and especially widowed members of the populations have increased importantly, while the number of cohabiting and "other" states have decreased (the majority of children belong to the latter state). Evidently, these changes are related to the ageing of the population. Following these changes, the numbers of single and double households and the number of households without children have increased significantly, while the number of larger (3+ people) households and the number of households with children have declined.

The composition of the population has also changed with regard to possession of tourist commodities. In particular, the possession of accommodations with a permanent location has increased, while the number of people with an accommodation without a permanent place has decreased. In addition the possession of cars has decreased slightly. The possession of skis, finally, has not changed significantly.

It is emphasised that due to the use of static ageing, these changes represent the *present* relationship between age and other socio-economic and -demographic variables. In other words, the impact analyses excludes cohort effects in the relationship between age and other socio-economic and –demographic variables. In the present demonstration project, for instance, the possession of one or more cars in the household decreases slightly (from 82.03% in 1998 to 81.22% in 2020), while future elderly are expected to have higher levels of car possession (Blom & Sahebdien, undated). In addition, it has been argued that *MERLIN* excluded cohort effects with regard to tourist choices. These limitations should always be kept in mind while interpreting the impact of the proposed changes.

OBSERVED IMPACT AND INTERPRETATION OF THESE CHANGES

Under the assumption of a constant size of the population, the proposed composition for the year 2020 will result in small but significant decreases in both the number of day-trips (-1.16%) and the number of holidays (-0.75%). The overall number of days not allocated to tourist trips (p.p.p.y.), however, will not change significantly, because the average duration of holidays increases significantly from 9.18 days at present to 9.28 days in 2020 (see Table A6.7). This shift towards longer holidays is also noticeable in the participation choices (Table A6.6): the number of short breaks decreases (-3.18%) while the number of medium long holidays increases (+1.34%), the number of long holidays decreases (-2.25%), while the number of extra long holidays increases dramatically (+15.66%). These observations correspond to the model parameters discussed in chapter 7, where it was concluded that, relative to the oldest age group, young adults (16-30 years) and adults (31-55 years) are likely to allocate more time to day-trips and short breaks and less time to longer (9+ days) holidays.

The participation choices of children (0-15 years) are dominated by the decline in the absolute number of children (under the assumption of a constant size of the population) because both the number of children without tourist trips (-4.14%) and the number of children pursuing holidays (-4.25%) decrease approximately 4%, while the average numbers of holidays (of various duration) per person do not change.

For adults, on the other hand, the proposed ageing of the population will increase the number of adults who do not pursue tourist trips at all (+12.65%) and the number of adults who only make day-trips (+2.52%), while the number of adults who pursue both day-trips and holidays will decrease (-1.91%). The time allocated to tourist trips by the adults who *do* pursue tourist trips, however, increases in all participation groups: the average number of day-trips of those who only make day-trips increases by 2.52%, while the number of days *not* allocated to tourist trips decreases significantly in the participation groups with holidays (-0.10% and -0.30% respectively). Based on these results it can be concluded that the increased share of the elderly will result in lower levels of participation, but once people have decided to pursue tourist trips, they are inclined to allocate more time to tourist trips. These patterns are probably explained best by the fact that a certain proportion of the elderly is incapable to travel due to health issues, while those that are still healthy have the time (and money?) to pursue tourist trips rather frequently.

As expected, the absolute number and share of holidays with children will decrease dramatically due to the greying of the population (-7.80% for holidays with schoolchildren and -7.61% for holidays with other children), while the number of holidays with adults only increases (4.43%). Strikingly, the increase in the number of single households will not result in a significant demand for single holidays. Apparently, the new and relatively old single households are not interested in travelling independently. The increased number of “unknown” travel parties (+1.58%), finally, should probably be attributed to the increased number of households with tourist accommodation with a fixed location and/or the higher travelling frequencies of those who pursue tourist trips.

Another interesting development resulting from the greying of the population is the relief of the traditional peak periods such as the summer season (-3.19% in the first, and –3.58% in the second part) and spring school holiday periods (-2.85%) and the concurrent increased demand for holidays in the off-peak periods (+3.35% in the winter outside school holiday periods; +1.62% in the spring outside school holiday periods; and +3.17% in the autumn outside school holiday periods). Also, congestion and other traffic problems will diminish to some extent because fewer people will depart on Fridays and Saturdays (-1.72%). These developments allow providers of tourist products and services to extend the tourist season and optimise their occupancy rates.

The greying of the population will also improve the position of domestic destinations relative to destination abroad, because, despite the decreased overall number of holidays, no significant changes occur for domestic destinations, whereas several neighbouring countries such as France (-2.48%) and Belgium and Luxembourg (-4.00%) become less popular. In terms of expenditures, however, a contrasting development occurs: the average expenditures on domestic holidays remain constant, while the average expenditures on foreign holidays in general (+1.69%), and long holidays abroad in particular (+1.56%), increase. Overall, the expenditures per holiday increase by +1.27%, partially due to the increased expenditures on (foreign) holidays and probably also due to the increased duration per holiday (given the serious underestimation of the *MERLIN*-system with regard to the average expenditures for (extra) long holiday (see section 12.3.2), the increase in expenditures on holidays may even be more pronounced than the present figures suggest).

In terms of accommodation choices, the analyses show that both the privately owned non-permanent (-2.84%) and the “other” (-2.55%) accommodations will become less popular, while holidays with/at privately owned accommodations with a permanent location will increase (+2.98%). The latter result is probably explained best by the increase in the number of households owning such an accommodation.

With regard to transport mode choices, finally, the analyses show that the reduced number of holidays mainly effects the use of cars (-1.10%). It has been noted before, however, that the number of households owning one or more cars is likely to increase in the future, which would violate the assumptions underlying the present analyses and compensate the observed effect at least partially.

Based on the above discussion, it can be concluded that, under the assumption of a constant size of the population and excluding cohort effects, the greying of the population will have a significant impact on the demand for tourist trips. First and foremost, the increased share of elderly will result in lower levels of participation, while at the same time the trip frequencies of those who do pursue tourist trips will increase. Overall, the demand for day-trips and holidays in general will decrease slightly, but the demand for longer holidays will increase. Furthermore, the share of holidays with children will decrease, while the holidays with adults only will increase.

The greying of the population also offers some interesting opportunities for the providers of tourist products and services to extend the tourist season and to increase their occupancy rates because the pressure on the traditional peak periods will diminish. Also, the Netherlands can improve its positions as a holiday destination for the Dutch people because the analyses show that the overall decrease in the annual number of holidays will mainly effect neighbouring countries. On the other hand, the analyses also show that, on average, people will spend more on foreign holidays (per holiday), while the expenditures on domestic trips remain constant. This trend may be strengthened by the increased use of privately owned tourist accommodation with a fixed location because holidays to/with these accommodation types tend to be frequent, domestic and relatively cheap. Finally, the analyses show that the greying of the population will decrease the demand for car-based holiday. It should be noted, however, that the analyses did not take into account cohort effects such as the (expected) increased possession of cars among the “new” elderly.

12.6 Conclusion and discussion

This last chapter but one has brought together the conceptual and empirical building blocks that were advanced in the previous chapters. Following the final adjustments to the empirical input into the MERCIN-system, the validity and the use of the system were discussed.

With regard to the validity of the system it was shown that the stability of the simulation results and the performance of the MERCIN-system at both the aggregate and the individual level is reasonable on a large number of indicators. There are, however, some system irregularities that should be kept in mind while using the system. First, the homogeneity of the population, proved to be rather fragile, because splitting the population in two equal segments produces many significant differences between the two halves. When the entire population of simulation units is used, however, the ability of the system to reproduce simulation results is very good, even at a relatively low number of runs (N = 25).

Second, the ability of the *MERLIN*-system to reproduce the original data is substandard in some cases. In particular, the number of “unknown” travel parties is underestimated, which also causes a serious underrepresentation of the number of holidays in tourist accommodations with a permanent location that are owned by the traveller. Also, there are some biases in the reproduction of several destinations. Finally, although a large number of different modelling approaches have been tested, even the “best” model severely overestimates the expenditures on short breaks, while the average expenditures on extra long holidays were seriously underestimated. Notwithstanding these abnormalities, however, overall, the ability of the system to reproduce the original data at the aggregate level proves to be quite reasonable.

Third, the abilities of the *MERLIN*-system to reproduce the observed tourist choices at the individual level was tested by regressing the average predicted choices with regard to the number of holidays and day-trips, the destination choices, the duration choices and the expenditures across 200 runs onto the observed choices. Except for the number of extra long holidays and to a lesser extent also the average duration of, and expenditures on holidays, these relationships prove to be quite reasonable. Finally, although the relationship between the observed and the predicted values at the individual level is adequate, the part of the variance for each indicator that is explained by the *MERLIN*-system proves to be rather low.

The final two sections of this chapter discussed and demonstrated the options available in the *MERLIN*-system to assess the impact of several social and demographic developments on tourist trip patterns. It was shown, for instance, that increasing the number of free days for 10% of the full-time (30+ hours per week) working part of the population, will only result in a small increase in the number of holidays, while the number of day-trips remains constant. Also, under the assumption of a constant size of the population, the ageing of the population will result in an increased demand for longer holidays and for holidays outside the traditional peak periods, while demand for car-based holidays will decrease. Finally, it was shown that the doubt cast in the previous chapters on the ability to assess the effect of changes in the aggregate preferences for transport modes proved to be legitimate. Although the technical reasons for this deficiency are evident, the real reasons behind it are not clear.

13 Conclusion and Discussion

This final chapter concludes this thesis by summarising its objective, by discussing the major research contributions and potential limitations and by identifying potentially promising areas of future research.

13.1 Objectives of this thesis

To support the spatial planning of outdoor leisure facilities, many theories and models of preference and choice have been advanced in the past to assess the potential impacts of policy and planning decisions. The majority of these studies focus on the characteristics of destinations (or other 'attractions') and use so-called utility-based approaches to quantify the contributions of the identified characteristics to the overall utility of these destinations. Based on more recent publications, however, it has been argued that the above approaches do not longer suffice for a variety of reasons. First, they focus on single trips and capture only part of the travel decision. It is now generally recognised, for instance, that the travel decision-making process comprises a number of interrelated trip decisions or *facets*, and is better described as a *trip profile*, including decisions regarding the members of the travel party, timing, transport mode, accommodation, destination choice, and so on. In addition, the complex interaction between preferences, opportunities and constraints in the decision-making process has recently drawn the attention of the research community. Also, it is now believed that subsequent trips may affect each other. This has a potentially dramatic impact on the prediction of future demand for facilities. Finally, the assumption of utility-maximising behaviour has been challenged for various reasons. The so-called constraints-based approaches, for instance, argue that people are not free to choose the trips they desire because they have to take into account restrictions such as limited time and money budgets or coupling constraints. Alternatively, approaches based on modern psychological and physiological theory argue that it is not realistic to assume individuals to have complete information of all alternatives nor that they are able (and willing) to arrive at optimal solutions.

Consequently, the majority of models used in the outdoor leisure planning practice today lack the sophistication required to gain insight into individual tourist trip patterns, i.e. interrelated sets of trips, each of which is characterised by a trip profile, that is pursued by an individual during a particular period of time. The need for more complex models is increased due to the still growing market orientation of policy-makers and the industry alike. To really understand these complex processes, much richer models, composed of more variables, and representing more complex relationships among those variables, should be developed. This thesis therefore aimed to develop and test such a model. In particular, a *microsimulation model* taking into account many facets of outdoor leisure trips and mutual dependencies among these facets and among subsequent outdoor leisure trips has been advanced and tested. Microsimulation offers the opportunity to build more comprehensive models of leisure choices by combining behavioural hypotheses and relationships with regard to the facets (or subdecisions) that comprise this decision. Aggregate effects, then, are obtained by combining these individual simulations. The outdoor leisure trips are referred to as *tourist trips*, because tourism definitions are traditionally rooted in notions of economics, time use and distance travelled.

13.2 Contributions and limitations

The model under development is part of a decision support system, referred to as the *MERLIN*-system, which assists planners and policy makers in evaluating alternative scenarios and planning options. Given the structure of the *MERLIN*-system, this thesis contributes to the *literature on tourist decision-making* and to *planning in the field of recreation and tourism* at four levels. In addition, this thesis also introduces two *methodological innovations* to the tourism research literature. This final section discusses these contributions and their potential limitations, and identifies possible subjects for future research.

13.2.1 Contributions and limitations with regard to the literature on tourist decision-making and to tourism planning

As argued before, this thesis contributes to the literature on tourist decision-making and to planning in the field of (recreation and) tourism at four levels. First, insights are provided into the *relationship between individual and household characteristics and a large number of tourist choices* that are relevant to the research and planning communities. The system includes the following tourist choices: the annual number of day-trips and holidays, and with regard to the holidays, for each trip the choice of travel party, timing (season and the relation with school holiday periods), destination, accommodation, transport mode, date of departure and the choice of expenditures. The set of variables that are assumed to condition the tourist choices include individual and household characteristics such as age, gender, household income, educational level, various conditions related to the working situation, civil state, the presence of children in the household, various conditions related to the place of residence and the availability of various (tourist) commodities such as cars, tourist accommodations and skis. Evidently, future research could explore other choice and/or condition variables. Due to data restrictions, for instance, *MERLIN* does not include any information with regard to ethnicity while several studies indicate non-native Dutch to pursue different outdoor leisure activities and/or holidays (Schmeink & Ter Wolde, 1999). Moreover, the share of non-natives within the Dutch population is expected to increase significantly in the future, thus increasing the importance of including these conditions in tourism-related studies and models. This information should therefore be included in future surveys on day-trips and holidays.

The second contribution of the *MERLIN*-system to the tourism literature on decision-making and to tourism planning is the focus on the *relationships between the various facets of tourist trips*. Although the majority of tourist choices have been contemplated in the tourism literature either on their own or in relation to other facets, the sheer number of facets incorporated in a *quantitative* model is unprecedented. In addition, by allowing choices in earlier stages of the decision-making process to condition subsequent choices, the system takes into account the interactions between the various facets of tourist trips. A major limitation in this respect is that *MERLIN* assumes a sequential choice process with regard to trip profiling, i.e. the order in which the facets of each tourist trip profile are considered. Evidently, this order affects the results because previous choices frame the condition space of subsequent choices. Within the *MERLIN*-system, several solutions were advanced to (partially) avoid some of these problems. In particular, this applies to expenditures because expenditures on trips with higher profiling priorities were included as conditions for decision of trips yet to be planned. In addition, a possible solution was advanced for the relation between destination and mode choices. With regard to expenditures, the proposed solution appeared to work quite well, although the condition variables expressing the total planned amount of costs on holidays were never among the most important conditions. The limitation of this solution is, however, that it only captures the effects of the characteristics of previously considered tourist trips. It does not captures the possible effects of choices on subsequent facets of the tourist trip under consideration. Future research should therefore explore he possibilities of simultantously deriving the decision rules for multiple choice facets.

The proposed solution for the bi-directional relationship between destination and transport mode choices comprised the inclusion of condition variables in the destination choice models that express (on a scale from 0 to 100) the tourist's general propensities to select particular transport mode alternatives for holidays (i.e., car, aeroplane or "alternative"). These propensities are based on a utility-based model that incorporates the aggregate marginal distribution of the transport mode alternatives and several characteristics of the individual and his or her household. The propensities should allow the user to assess the effects of changes in the general preferences for transport modes (e.g., due to increasing costs) on tourist trip patterns by changing the aggregate marginal distribution over the modes. Based on several tests, however, it was concluded that the decision rules for destination choices did not "look ahead" as expected. The technical reason in terms of the models' structures is evident: the "propensity conditions" are hardly included in the condition parts of the DTs for destination and transport choices. The "reluctance" of these choice models to integrate the individual transport mode preferences in their condition structures, on the other hand, is less obvious.

Third, MERCIN describes the *relationship between the tourist trips pursued by one individual*. In this context, for instance, the system captures the conditions under which repetitive and/or variety seeking choice behaviour occurs by including characteristics of other tourist trips as conditions for subsequent choices. It was shown, for instance, that when people have already planned one or more holidays at more distant destinations, the probability of visiting another more distant destination decreases dramatically. In contrast, in the choice between domestic regions and tourist areas (and to some extent also in the choices between the neighbouring countries), repeat visitation patterns were often observed. In these cases, when one or more holidays had already been planned for a particular area, the probability of visiting that destination again increased. Analogous to the relationship between the facets of tourist trips, the relationships between subsequent tourist trips are limited by the fact that MERCIN assumes the *profiling priority* of a tourist trip within an individual's trip program to be fixed. These priorities are first based on the duration, and second, if the duration is identical, on the travel party. Although this profiling priority thus takes into account the possible differences between main holidays (that are often longer, more distant, more expensive and/or have more extensive planning horizons) and extra holidays, and also reflects the idea that people are constrained by other people's agenda's, these assumptions do reduce the complexity of the system. It would be an interesting subject for future research to explore other conditions that determine the planning priority of tourist trips (e.g., the timing of a trip or the similarity between tourist trip patterns in consecutive years) and how these conditions can be incorporated in comprehensive systems like MERCIN.

Another novelty with regard to the relationship between the tourist trips introduced by *MERLIN*, is that day-trips and holidays are considered together at the level of the individual. Although each of these types of trips has been studied in considerable detail, little was known about the interdependencies, if any, between day-trips, and holidays in the tourist decision-making process. Theoretically, however, it was hypothesised that day- and overnight trips should be related because any consumption decision necessarily involves some allocation of time and money - resources that are ultimately finite. Within the framework of this thesis, a pilot study was conducted and a general approach to modelling time allocation/trip generation was proposed. It was concluded that people who allocate more time to day-trips are more likely to allocate less time to longer overnight holidays and the other way around. Also, it was found that older people are more likely to pursue longer holidays whereas young adults are more likely to pursue day-trips and shorter holidays. A major limitation in this respect was the availability of data on day-trip and holiday behaviour at the level of the individual because surveys on day-trips and holidays in the Netherlands are conducted independently. The empirical models that constitute the building blocks of the *MERLIN*-system are based on the 1998 Continuous Vacation Research (i.e., the CVO-survey) supplemented with estimates (provided by the respondents) on annual day-trip frequencies. These data exclude, however, the exploration of possible relationships between the facets of day-trips and those of holidays. It would be interesting to identify these relationships, if any.

The final contribution to tourism literature on decision-making and tourism planning is that the *MERLIN*-system allows policy makers and suppliers to *evaluate the impact of various scenarios for the future in terms of entire annual tourist trip patterns*[43]. To this end, the system offers an user-friendly, menu-guided interface that allows the user to change the characteristics of (a segment of) the population and assess the impact of these changes on annual tourist trip patterns by statistically comparing the simulation results to the present situation. In this respect, the use of the system is evidently limited to “what has been put into the system”. In other words, “to what changes is the system sensitive?” There are three types of policy options available in the *MERLIN*-system, including:

1. Changes at the level of tourist trip programs that were obtained using stated preference techniques;

[43] It is acknowledged that the with regard to recreation and tourism planning, the contributions at the previous three levels are rather indirect. In these cases, planning is supported by increasing the general knowledge with regard to outdoor leisure (choice) behaviour. The contributions at this fourth level, however, are relevant to planners and policymakers more directly (and mainly at the national level).

2. Changes in the aggregate preferences for transport mode alternatives; and
3. Changes in the characteristics and/or composition of the population.

With regard to the second option, it has already been argued why these options should not be used for policy and/or scenario evaluations. With regard to the first policy options, 7 scenarios that were presented to the members of the CVO-panel are included. The third type of policy options offers the opportunity to assess the impact of changes in the characteristics and/or composition of the population. Using this option, for instance, it was demonstrated that, under the assumption of a constant population size, the ageing of the population will result in an increased demand for longer holidays and for holidays outside the traditional peak periods, while demand for car-based holidays will decrease.

The main limitation of *MERCIN* with regard to the evaluation of scenarios for the future is that the majority of the available options are related to the demand side of the market. The possibilities to assess the impact of changes on the supply side of the market are limited to the conditions of the “Student Public Transport Ticket” and the conditions of the “Benefit Hours Ticket” for trains. The impact of other changes on the supply side can only be assessed indirectly. In other words, the user has to “translate” these changes into the options available to the system. Assuming the number of people owning a tourist accommodation with a permanent location to decrease, for instance, could be an approach to assessing the impact of an increase in the price of seasonal places. Other supply side changes, such as the price, the quality and the capacity of the (various sectors of) the tourist-recreation product (TRP) are simply not available in the present version of the system. A possible approach might be to use longitudinal data, e.g. several consecutive years of the CVO-panel survey, including additional data (and/or assumptions) regarding the price, the quality and the capacity of the domestic and foreign TRP, and including these data as conditions in the empirical building blocks of the system. An additional advantage of longitudinal data would be the opportunity to identify the effects of conditions (preferences and constraints) that are permanent and distinguish them from structural variations across several years, lagged responses and “real” random effects due to, for instance, short-lived trends and hypes and other incidental factors.

Finally, the present version of the system conveniently assumes that the 1998 CVO-sample of respondents represents the entire population of interest. To aid scenario analysis and forecasting for planning applications, the simulation results should be scaled up to the entire Dutch population. Using scale parameters that represents the weight of each respondent within the Dutch population would achieve this. The application of these scale parameters, however, requires additional care with regard to the match between the base of these scale parameters and the variables that condition the tourist choices in the system because possible mismatches may increase existing biases.

13.2.2 Methodological contributions and limitations

In addition to the above-mentioned contributions, this thesis also introduced two *methodological innovations* to the tourism research community. First, a *method for deriving choice heuristics* (or *decision rules*) *from empirical data* for each stage of the decision-making process was advanced. Although research on tree-induction algorithms has been appearing in several other disciplines, the topic is still largely unexplored in tourism research The majority of the above mentioned tourist choices are described using so-called probabilistic decision rules (IF < condition > THEN < action> statements) established in probabilistic Decision Tables (DTs) that indicate under which conditions tourists will take a particular decision with a certain probability. The focus on choice heuristics is relatively rare in tourism research. Although it has been suggested by several authors that travellers use heuristics, or choice rules, to set priorities for their choice decisions (King & Woodside, 2001; Bervaes *et al.*, 1996; Bronner & De Hoog, 1985), operational rule-based models of tourist and recreation behaviour are rare. An exception includes two studies by Law and Au, who, based on the rough set theory, induced two groups of decision rules to predict tourist expenditures (Law & Au, 2000; Au & Law, 2000). However, their choice of method implied that they could not generate exhaustive sets of rules. In contrast, by inducing decision rules using an algorithm based on the traditional Chi-square Automatic Interaction Detection (CHAID) analysis, this thesis introduces an approach to research on tourist choice behaviour that generates sets of choice heuristics that are exclusive, exhaustive and complete. In other words, the CHAID-based algorithm generates rule bases that allow one to uniquely classify every possible tourist choice situation. Moreover, the statistical properties of the generated DTs proved to be satisfactory.

Compared to the more traditional utility-based approaches, the proposed rule-based approach to tourist decision-making relieves some of the limitations because it offers more flexibility in modelling tourist choice behaviour. In particular with regard to the model specification, rule-based models do not assume an a priori functional form for the model, nor do they require the explicit inclusion of interactions (i.e. consciously deliberated relationships) between condition states. There are, however, two obvious difficulties in deriving and using rule-based models. One of the main disadvantages is that the sets of decision rules can become rather large and it may sometimes be difficult to represent them economically. On the other hand, utility-based models may suffer from similar problems and can also become very complex. The advantages and disadvantages with regard to the communication and interpretation of the results of both approaches should be compared.

Second, at present, little is known about the possibilities of the various rule-inducing algorithms. The CHAID-based algorithm was selected because of its greater sensitivity to the whole response distribution, which is favourable in the light of probabilistic decision rules. It was acknowledged, however, that this comes at the cost of more elaborate sets of decision rules and that, given the present state of knowledge regarding the application of rule-induction systems in the social sciences, the choice in favour of any algorithm is at least partially arbitrary. With regard to the CHAID-based algorithm, it has been argued, for instance, that future research should focus on the optimal settings for this algorithm (Van Middelkoop *et al.*, 2000; see Van Middelkoop *et al.* (2001) for an extended version). To meet this argument, sensitivity analyses were conducted for each DT in the MERCIN-system. In addition, however, the characteristics and effects of other rule-inducing algorithms such as C4.5 and CART should also be examined in the context of modelling tourist choice behaviour.

The second methodological contribution to the field of tourism is that MERCIN is *one of the first operational models* that incorporates a large number of choice facets, the interactions between these facets and the interactions between tourist trips that are separated in time. Although the complexity of the tourist decision-making process has been acknowledged for some time now, few attempts have been made to model this process quantitatively appreciating its full complexity. A review of simulation systems in the field for recreation and tourism, for instance, produced a list of only five examples world-wide. And even these examples differed significantly in the level of complexity and the number of interactions incorporated.

Notwithstanding the apparent advantages, the decision to use a simulation approach inherently brings about several difficulties due to the costs and the time involved in developing, testing and using a simulation system. With regard to testing or validating the system, MERCIN is clearly a *model- or theory based simulation system* since the performance of the empirical building blocks as well as the system as a whole have been assessed. With regard to the validity of the whole system it was shown that the stability of the simulation results and the performance of the system on both the aggregate and the individual level was quite reasonable on a large number of indicators. There were, however, some system irregularities that should be kept in mind while using the system. Several improvements have already been incorporated to reduce the extent of these irregularities, but future research should continue to improve the performance of the system with regard to these indicators.

To conclude, *MERLIN* represents a first attempt to model multiple facets of the tourist decision-making process in an integrative fashion. The results provide support for the more recent publications that argue that tourist-decision making is a complex process, comprising many facets and interactions between these facets. Although the system already performs reasonably well, several other alternatives exist for some of the operational decisions made. Particular aspects of the system can therefore be further elaborated in future research. The results of this thesis, however, strongly suggest that any attempt to model tourist choices that disregards the inherent complexity is sensitive to significant biases and may exclude important aspects of tourist choice behaviour or even result in misleading policy guidelines.

References

Arentze, T.A, F. Hofman & H.J.P. Timmermans, (2001), Deriving Rules from Activity Diary Data: A Learning Algorithm and Results of Computer Experiments. To appear in *Geographical Systems*.

Arentze, T.A., F. Hofman, H. van Mourik, H.J.P. Timmermans & G. Wets (2000), Using Decision Tree Induction Systems for Modelling Space-Time Behaviour. *Geographical Analysis,* 32, 52-72.

Arentze, T.A., G.L. Lucardie, H. Oppewal & H.J.P. Timmermans (1996), A Functional Decision Table Based Approach to Multi-Attribute Decision Making. *3rd International Conference on Retailing and Consumer Services Sciences*. Telfs/Buchen, Austria.

Arentze, T.A. & H.J.P. Timmermans, (2000a), Decision Rule Formalisms. In: T.A. Arentze & H.J.P. Timmermans (Eds.), *ALBATROSS: A Learning Based Transportation Oriented Simulation System* (pp. 108-135), Eindhoven: EIRASS.

Arentze, T.A. & H.J.P. Timmermans (2000b), The *ALBATROSS* system. In: T.A. Arentze & H.J.P. Timmermans (Eds.), *ALBATROSS: A Learning Based Transportation Oriented Simulation System* (pp. 81-107), Eindhoven: EIRASS.

Arentze, T.A. & H.J.P. Timmermans (2000c), Conceptual framework. In: T.A. Arentze & H.J.P. Timmermans (Eds.), *ALBATROSS: A Learning Based Transportation Oriented Simulation System* (pp. 71-80), Eindhoven: EIRASS.

Au, N. & R. Law (2000), The Application of Rough Sets to Sightseeing Expenditures. *Journal of Travel Research*, 39, 70-77.

Bargeman, B., C.H. Joh & H.J.P. Timmermans (1999), Correlates of Tourist Vacation Behavior: A Combination of CHAID and Loglinear Analysis. *Tourism Analysis*, 4, 83-93.

Baxter, M. & G. Ewing, (1981), Models of Recreational Trip Distribution. *Regional Studies*, 15, 327-344.

Beckers, T.A.M. (1983), *Planning voor Vrijheid. Een Historisch-Sociologische Studie van de Overheidsinterventie in Recreatie en Vrije Tijd*. Ph.D.-thesis, Mededelingen van de vakgroepen voor sociologie nr. 8, Wageningen: Landbouwhogeschool Wageningen.

Beckers, T.A.M. (1995), Tourism Planning in the Netherlands: Back to the Future. *Tijdschrift voor Economische en Sociale Geografie*, 86, 93-97.

Ben-Akiva, M. & S.R. Lerman (1985), *Discrete Choice Analysis. Theory and Application to Travel Demand*. London: The MIT Press.

Ben-Akiva, M. & T. Morikawa (1990), Estimation of Travel Demand Models from

Multiple Data Sources. In: M. Koshi (Ed.), *Transportation and Traffic Theory,* (pp. 461-476), New York: Elsevier.

Bervaes, J.A.C.M., H.J.J. Kroon, G.F.P. Martakis & D.C. van der Werf (1996), *Een Model voor het Gebruik van de Groene Ruimte in Stadslandschappen* (Fase I), Wageningen: Instituut voor Bos- en Natuuronderzoek IBN-DLO.

Bhat, C.R. & F.S. Koppelman (1993), A Conceptual Framework of Individual Activity Program Generation. *Transportation Research A*, 27A, 433-446.

Bhat, C.R. & R. Misra (1999), Discretionary Activity Time Allocation of Individuals between In-Home and Out-of-Home and Between Weekdays and Weekends. *Transportation*, 26, 193-209.

Biggs, D., B. de Ville & E. Suen (1991), A Method of Choosing Multiway Partitions for Classification and Decision Trees. *Journal of Applied Statistics*, 18, 49-62.

Bishop, I. & H.R. Gimblett (2000), Management of Recreational Areas: GIS, Autonomous Agents and Virtual Reality. *Environment and Planning B*, 27, 423-435.

Blom, P. & A. Sahebdien (undated), *7 Trends. Mobiliteit in Veranderend Nederland*. Rotterdam: Adviesdienst Verkeer en Vervoer, Afdeling Maatschappelijke Verkenningen en Beleidsanalyse (source: http://www.minvenw.nl; December 2000).

Bonn, M.A., H.L. Furr & M. Uysal (1992), Seasonal Variation of Coastal Resort Visitors: Hilton Head Island. *Journal of Travel Research*, 30, 50-56.

Borgers, A.W.J. & H.J.P. Timmermans (1996), *Modelling System for the Toermodel*. Eindhoven: EIRASS.

Borgers, A.W.J., R.E.C.M. van der Heijden & H.J.P. Timmermans (1987), A Variety Seeking Model of Spatial Choice-Behaviour. *Environment and Planning A*, 21, 1037-1048.

Brazier, F.M.T., B.M. Dunin-Keplicz, N.R. Jennings & J. Treur (1997), DESIRE: Modelling Multi-Agent Systems in a Compositional Formal Framework. *International Journal of Cooperative Information Systems*, 6, 67-94.

Breiman, L., R.A. Friedman, R.A. Olshen & C.J. Stone (1984), *Classification and Regression Trees*. Belmont (CA): Wadsworth.

Bronner, A. & R. de Hoog, (1985), A Recipe for Mixing Decision Ingredients. *European Research*, 109-115.

Brouwer, N., G. Beekers & H. Verheijden (1994), *Vrije Tijd en Vakanties: Een Statistische Relatie?* Amsterdam: SEO.

Brouwer, E., H. van Dijk & E. Oldenboom (1995), *Dagtochtmodel.* Amsterdam: SEO.

Bunch, D.S. (1991), Estimability in the Multinomial Probit Model. *Transportation*

Research-B, 25B, 1-12.

CBS (1992), Dagrecreatie 1990/'91. Voorburg/Heerlen: CBS.

CBS (1996), Vakanties van Nederlanders 1995. Voorburg/Heerlen: CBS.

CBS (1997), Dagrecreatie 1995/'96. Voorburg/Heerlen: CBS.

CBS (2001), CBS-Statline, Source: http://www.cbs.nl/nl/statline (April 2001)

Cesario, F.J. (1973), A Generalised Trip Distribution Model. *Journal of Regional Science*, 13, 233-248.

Cesario, F.J. (1975), A Simulation Approach to Outdoor Recreation Planning. *Journal of Leisure Research*, 7, 38-52.

Chen, J. S. (1997), The Tourist's Cognitive Decision Making Model. *Revue de Tourisme*, 4, 4-8.

Cohen, E.A. (1979), A Phenomenology of Tourist Experiences. *The Journal of British Sociological Association*, 13(2), May 1979.

Crawford, D.W. & G. Godbey (1987), Reconceptualizing Barriers to Family Leisure. *Leisure Sciences*, 9, 119-127.

Crompton, J.L. (1992), Structure of Vacation Destination Choice Sets. *Annals of Tourism Research*, 19, 420-434.

Crouch, G.I. & J.J. Louviere (2000), A Review of Choice Modelling Research in Tourism, Hospitality and Leisure. *Tourism Analysis*, 5, 97-104.

Dadgostar, B. & R.M. Isotalo (1992), Factors Affecting Time Spent by Near-Home Tourists in City Destinations. *Journal of Travel Research*, 31, 34-39.

Daganzo, C. (1979), *Multinomial Probit. The Theory and Its Applications to Demand Forecasting*: Academic Press.

Davies, F., M. Goode, J.A. Mazanec & L. Mountinho (1999), LISREL and Neural Network Modelling: Two Comparason Studies. *Journal of Retailing and Consumer Services*, 6, 249-261.

Davis, H.L. & B.P. Rigaux (1974), Perception of Marital Roles in Decision Processes. *Journal of Consumer Research*, 1, 51-62.

Dellaert, B.C.G. (1995), *Conjoint Choice Models for Urban Tourism Planning and Marketing*. Ph.D.-thesis, Eindhoven University of Technology.

Dellaert, B.C.G., A.W.J. Borgers & H.J.P. Timmermans (1997), Conjoint Models of Tourist Portfolio Choice: Theory and Illustration. *Leisure Sciences*, 19, 31-58.

Dellaert, B.C.G., D.F. Ettema & C. Lindh (1998a), Multi-Faceted Tourist Travel Decisions: A Constraint-Based Conceptual Framework to Describe Tourists' Sequential Choices of Travel Components. *Tourism Management*, 19, 313-320.

Dellaert, B.C.G., M. Prodigalidad & J.J. Louviere (1998b), Using Conjoint Analysis to Study Family Travel Preference Structures: A Comparison of Day-Trips and 1-week holidays. *Tourism Analysis*, 2, 67-75.

Dietvorst, A.G.J. (1987), *Beeldvorming en Keuze van Vakantiegebieden in Nederland.* Research report, Nijmegen: Werkgroep Recreatie en Toerisme, Geografisch en planologisch instituut, Katholieke Universiteit Nijmegen.

Dietvorst, A.G.J. (1989), *Complexen en Netwerken: Hun Betekenis voor de Toeristisch-Recreatieve Sector*. Wageningen: Agricultural University.

Dietvorst, A.G.J. (1993a), Planning for Outdoor Recreation and Tourism in the Netherlands. In: H.N. van Lier & P. D. Taylor (Eds.), *New Challenges in Recreation and Tourism Planning* (pp. 69-86), Amsterdam: Elsevier.

Dietvorst, A.G.J. (1993b), Planning for Tourism and Recreation: a Market-Oriented Approach. In: H.N. van Lier & P.D. Taylor (Eds.), *New Challenges in Recreation and Tourism Planning* (pp. 87-124), Amsterdam: Elsevier.

Dietvorst, A.G.J. (1993c), *Tourist Recreation Development and Spatial Transformations*. Wageningen: Agricultural University.

Dietvorst, A.G.J. (1995), Tourist Behaviour and the Importance of Time-Space Analysis. In: G.J. Asworth & A.G.J. Dietvorst (Ed.), *Tourism and Spatial Transformations: Implications for Policy and Planning* (pp. 163-181).

Dirven, R.J.A.G., W.A.M. van Hal, W.L.G. Sliepen & L.J. Wijker (1998), *Vakantiekeuzen van Nederlanders.* Research report, Breda/Voorschoten: NRIT/Toerisme & Recreatie AVN.

Eliasson, J. & L.G. Mattsson (1998), Modelling Long-Period Trip Patterns with Time and Money Constraints - the TILT Land Use Transportation Model. *WCTR Conference*. Antwerp.

Enneking, T. (1979), *Het Keuzebeslissingsproces met Betrekking tot de Vakantie binnen het Gezin*. Research report.

Ettema, D.F. (1996), *Activity-Based Travel Demand Modeling*. Ph.D.-thesis, Eindhoven University of Technology.

Ettema, D.F. & H.J.P. Timmermans (1997), Theories and Models of Activity Patterns. In: D.F. Ettema & H.J.P. Timmermans (Eds.), *Activity-based Approaches to Travel Analysis* (pp. 1-36), Amsterdam: Elsevier.

Feather, P., D. Hellerstein & T. Tomasi (1995), A Discrete-Count Model of Recreation Demand. *Journal Environmental Economics & Management*, 29, 214-227.

Fesenmaier, D.R. (1985), Modelling Variation in Destination Patronage for Outdoor Recreation Activity. *Journal of Travel Research*, 23, 17-23.

Fisher, M.M., G. Haag, M. Sonis & W. Weidlich (1990), Account of Different Views in Dynamic Choice Processes. In: M.M. Fisher, P. Nijkamp & Y.Y. Papageorgiou (Eds.), *Spatial Choices and Processes* (pp. 17-47), Amsterdam

Elsevier Science Publishers B.V. (North Holland).

Fishwick, P.A. (1995), *Computer Simulation: The Art and Science of Digital World Construction*. Gainsville: Computer & Information Science and Engineering Department, University of Florida.

Frechtling, D.C. (1997), Employing Monte Carlo Simulations to Forecast International Inbound Visitor Expenditures. In: *28th Annual TTRA Conference Proceedings* (pp. 146-155). Norfolk, Virginia, USA.

Ganglmair, A. & B. Wooliscrof (2001), K-Means vs. Topoplogy Representing Networks: Increasing Segment Stability in Tourism. In: J.A. Mazanec, G.I. Crouch, J.R. Brent Ritchie & A.G. Woodside (Eds.), *Consumer Psychology of Tourism, Hospitality and Leisure Volume Two* (pp. 231-242): CABI Publishing.

Gimblett, H.R. (1996), *Simulating Recreation Behaviour in Complex Wilderness Landscapes Using Spatially-Explicit Autonomous Agents*. PhD.-thesis, University of Melbourne.

Gimblett, H.R. & R.M. Itami (2000), RBSim - Recreation Behavior Simulator. Source: http://nexus.srnr.arizona.edu/~gimblett/rbsim.html

Glyptis, S. (1989), Life Styles and Leisure Patterns: Methodological Approaches. In: S. Glyptis & W. Tokarski (eds), *Life Styles: Theories, Concepts, Methods and Results of Life Style Research in International Perspective*. Prague: Academy of Sciences.

Goldberg, D.E. (1989), *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading: Addison-Wesley Publishing Company.

Greene, D.P. & S.F. Smith (1987), A Genetic System for Learning Models of Consumer Choice. In: J.J. Grefenstette (Ed.), *Second International Conference on Genetic Algorithms* (pp. 217-233), Massachusetts Institute of Technology, Cambridge, MA: Lawrence Erlbaum Associates.

Hägerstrand, T. (1970), What About People in Regional Science? *Papers of the Regional Science Association*, XXIV, 7-21.

Hanemann, M. (1984), Discrete/Continuous Models of Consumer Demand. *Econometrica*, 52, 535-561.

Hansen, F. (1976), Psychological Theories of Consumer Choice. *Journal of Consumer Research*, 3, 117-141.

Harding, A. (1996), Chapter 1: Introduction and Overview. In: A. Harding (Ed.), *Microsimulation and Public Policy. Selected Papers from the IARIW Special Conference on Microsimulation and Public Policy, Canberra, 5-9 Dec. 1993* (pp. 1-20). Amsterdam: Elsevier North Holland.

Hausman, J.A., G.K. Leonard & D. McFadden (1995), A Utility-Consistent, Combined Discrete Choice and Count-Data Model. *Journal Public Economics*, 56, 1-30.

Hiebeler, D. (1994), *The Swarm Simulation System and Individual-Based Modeling*. Santa Fe Institute working paper 94-12-065. Source: http://cam.cornell.edu/~hiebeler/swarm-paper.html

Holbrook, M.B. & D.R. Lehmann (1981), Allocating Discretionary Time: Complementarity Among Activities. *Journal of Consumer Research*, 7, 395-406.

Holecek, D.F. (1993), Trends in World-Wide Tourism. In: H.N. van Lier & P.D. Taylor (Eds.), *New Challenges in Recreation and Tourism Planning* (pp. 17-34), Amsterdam: Elsevier.

Inskeep, E. (1998), Tourism Planning: An Emerging Specialization. *APA Journal*, 360-372.

Iso-Ahola, S.E. (1982), Toward a Social Psychological Theory of Tourism Motivation: A Rejoinder. *Annals of Tourism Research*, 12, 256-262.

Iso-Ahola, S.E. (1984), Social Psychological Foundations of Leisure and Resultant Implications for Leisure Counselling. In: E.T. Dowd (Ed.), *Leisure Counselling: Concepts and Applications* (pp. 97-125), Springfield IL: Charles C. Thomas.

Jackson, E.L. (1988), Leisure Constraints: A Survey of Past Research. *Leisure Sciences*, 10, 203-215.

Jackson, E.L. (1990), Variations in the Desire to Begin a Leisure Activity: Evidence of Antecedent Constraints? *Jounral of Leisure Research*, 22, 55-70.

Jackson, E.L., D.W. Crawford & G. Godbey (1993), Negotiation of Leisure Constraints. *Leisure Sciences*, 15, 1-11.

Jansen-Verbeke, M.C. (1988), *Leisure, Recreation and Tourism in Inner Cities: Explorative Case-Studies*. Amsterdam/Nijmegen: Nederlandse Geografische Studies.

Jansen-Verbeke, M.C. & R. Spee (1995), *25 jaar vakantie van de Nederlander*. Rotterdam: Centrum voor Toerisme Management, Erasmus Universiteit.

Jeng, J.M. (1995), *Back-Propagation Neural Network Approach to Discrete Choice Modelling*. M.Sc.-thesis, University of Illinois at Urbana-Champaign.

Jeng, J.M. & D.R. Fesenmaier (1997), Facets of the Complex Trip Decision-Making Process. In: *28th Annual Conference of the Travel and Tourism Research Association* (pp. 31-41), Norfolk, Virginia USA.

Jenkins, R.L. (1978), Family Vacation Decision-making. *Journal of Travel Research*, 16, 2-7.

Johnson, N.L. & S. Kotz (1970), *Distributions in Statistics: Continuous Univariate Distributions 1*. Boston: Houghton, Mifflin Company.

Kass, G.V. (1980), An Exploratory Technique for Investigating Large Quantities of Categorical Data. *Applied Statistics*, 29, 119-127.

Kay, T. & G. Jackson (1991), Leisure Despite Constraints: The Impact of Leisure Constraints on Leisure Participation. *Journal of Leisure Research*, 23, 301-313.

Kemperman, A.D.A.M. (2000), *Temporal Aspects of Theme Park Choice Behaviour. Modeling Variety Seeking, Seasonality and Diversification to Support Theme Park Planning*. Ph.D.-thesis, Eindhoven University of Technology.

Kemperman, A.D.A.M., A.W.J. Borgers & H.J.P. Timmermans (2000), Consumer Choice of Theme Parks: A Conjoint Choice Model of Seasonality Effects and Variety Seeking Behavior. *Leisure Sciences*, 22, 1-18.

King, R.L. & A.G. Woodside (2001), Qualitative Comparative Analysis of Travel and Tourism Purchase-Consumption Systems. In: J.A. Mazanec, G.I. Crouch, J.R. Brent Ritchie & A.G. Woodside (Eds.), *Consumer Psychology of Tourism, Hospitality and Leisure. Volume Two* (pp. 87-106): CABI Publishing.

Kitamura, R., T. Yamamoto, S. Fujii & S. Sampath (1996), A Discrete-Continuous Analysis of Time Allocation to Two Types of Discretionary Activities Which Accounts for Unobserved Heterogeneity. In: J.B. Lesort (Ed.), *Transportation and Traffic Theory* (pp. 431-453), Oxford: Elsevier.

Kolsté, K. (1997), *Simulatie van Recreatief Tijd-Ruimte Gedrag; Evaluatie van het Programma RBSim*. MSc.-thesis, Agricultural University Wageningen.

Kraan, M.E. (1996), *Time to Travel?* Ph.D.-thesis, University of Twente.

Lancaster, K.J. (1966), A New Approach to Consumer Theory. *Journal of Political Economy*, 74, 132-157.

Law, R. & N. Au (1999), A Neural Network Model to Forcast Japanese Demand for Travel to Hong Kong. *Tourism Management*, 20, 89-97.

Law, R. & N. Au (2000), Relationship Modeling in Tourism Shopping: A Decision Rules Induction Approach. *Tourism Management*, 21, 241-249.

Lawson, R. (1991), Patterns of Tourist Expenditure and Types of Vacation across the Family Life Cycle. *Journal of Travel Research*, 29, 12-18.

Leonardi, G. & Y.Y. Papageorgiou (1992), Conceptual Foundations of Spatial Choice Models. *Environment and Planning A*, 24, 1393-1408.

Levine, R.L. & W. Lodwick (1983a), Continuous Simulation Methodology: A System Dynamics Approach to Planning, Forecasting, and Analysis of Recreational Usage. In: S.R. Lieber & D.R. Fesenmaier (Eds.), *Recreation Planning and Management* (pp. 162-186), London: E. & F.N. Spon Ltd.

Levine, R. L. & W. Lodwick (1983b), Introduction to Dynamic Techniques for Forecasting Recreation Behaviour. In: S.R. Lieber & D.R. Fesenmaier (Eds.), *Recreation Planning and Management* (pp. 141-161), London: E. & F.N. Spon

Ltd.
Lieber, S.R., & D.R. Fesenmaier (1984), Modelling Recreation Choice: A Case Study of Management Alternatives in Chicago. *Regional Studies*, 18, 31-43.
Lindh, C., B.C.G. Dellaert & D.F. Ettema (1995), Longer Term Activity Scheduling of Overnight Long Distance Trips: A Longitudinal Telephone Survey. In: *Workshop on Activity Based Approaches: Activity Scheduling and the Analysis of Activity Patterns*. Eindhoven, The Netherlands.
Louviere, J.J. & H.J.P Timmermans (1990a), A Review of Recent Advances in Decompositional Preference and Choice Models. *Tijdschrift voor Economische en Sociale Geografie*, 81, 214-224.
Louviere, J.J. & H.J.P. Timmermans (1990b), Stated Preference and Choice Models Applied to Recreation Research: A Review. *Leisure Sciences*, 12, 9-32.
Luce, R.D. (1959), *Individual Choice Behaviour: A Theoretical Analysis*. New York: Wiley.
Magidson, J. (1995), The CHAID Approach to Segmentation Modeling: CHI-Squared Automatic Interaction Detection. In: R.P. Bagotti (Ed.), *Advanced Methods of Marketing Research* (pp. 118-159): Basil Blackwell Ltd.
Mannell, R.C. & S.E. Iso-Ahola (1987), Psychological Nature of Leisure and Tourism Experience. *Annals of Tourism Research*, 14, 314-331.
Mansfeld, Y. (1992), From Motivation to Actual Travel. *Annals of Tourism Research*, 19, 399-419.
Mazanec, J.A. (1992), Classifying Tourists into Market Segments: A Neural Network Approach. *Journal of Travel and Tourism Marketing*, 1, 39-59.
Mazanec, J.A. (1999), Simultaneous Positioning and Segmentation Analysis with Topologically Ordered Feature Maps: A Tour Operator Example. *Journal of Retailing and Consumer Services*, 6, 219-235.
Mazanec, J.A. & L. Moutinho (1999), Editorial. Why is it Timely to Publish a JRCS Special Issue on Neural Networks. *Journal of Retailing and Consumer Services*, 6, 183-184.
McFadden, D. (1974), Conditional Logit Analysis of Qualitative Choice Behavior. In: P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105-142.), New York and London: Academic Press.
McFadden, D. (1978), Modeling the Choice of Residential Location. In: A. Karlqvist *et al*. (Eds.), *Spatial Interaction Theory and Planning Models* (pp. 75-96), Amsterdam: North Holland.
Meloni, I., E. Cherchi & L. Guala, (1998), *The Impact of Improvement in Transportation Supply on the Participation in Out-of-Home Activities*. Cagliari, Italy: Centre for Research in Transport Modelling.

Merz, J. (1991), Microsimulation - A Survey of Principles, Developments and Applications. *International Journal of Forecasting*, 7, 77-104.

Minar, N., R. Burkhart, C. Langton & M. Askenazi (1996), *The Swarm Simulation System: A Toolkit for Building Multi-Agent Simulations*.

Ministry of LNV (1997), *Recreatieschets 2020*. Den Haag: Department of Green Areas and Recreation, Ministry of Agriculture, Nature Conservation and Fisheries.

Mitra, A. & S. Lankford (1999), *Research Methods in Park, Recreation and Leisure Services*. Champaign, IL: Sagamore Publishing.

Morey, E.R., W.D. Shaw & R.D. Rowe (1991), A Discrete Choice Model of Recreational Participation, Site Choice, and Activity Valuation when Complete Trip Data are not Available. *Journal of Environmental Economics and Management*, 20, 181-201.

Morley, C.L. (1994), Experimental Destination Choice Analysis. *Annals of Tourism Research*, 21, 780-791.

MuConsult (1997), *Probleemveld Sociaal-Recreatief Verkeer. Hoofdrapport & Bijlagen*. MuConsult, in commission of the Projectbureau Integrale Verkeers- en Vervoersstudies.

NBT (1989), *Onderzoek naar het Maken van Dagtochten onder Nederlanders*. Leidschendam: NBT.

Oliver, J.R. (1994), Finding Decision Rules with Genetic Algorithms. *AI Expert*, 9, 33-39.

Olshavsky, R.W. (1979), Task Complexity and Contingent Processing in Decision Making: A Replication and Extension. *Organizational Behavior and Human Performance*, 24, 300-316.

Oppermann, M. (1995), Travel Life Cycle. *Annals of Tourism Research*, 22, 535-552.

Oppermann, M. (1998), Travel Horizon: a Valuable Analysis Tool? *Tourism Management*, 19, 321-329.

Orcutt, G.H. (1957), A New Type of Socio-Economic Systems. *Review of Economics and Statistics*, 58, 773-797.

O'Sullivan, D. & M. Haklay (2000), Agent-Based Models and Individualism: Is the World Agent-Based? *Environment and Planning A*, 32, 1409-1425.

Palvia, S.C. & S.R. Gordon (1992), Tables, Trees and Formulas in Decision Analysis. *Communications of the ACM*, 35, 104-113.

Parsons, G.R. & M.J. Kealy (1995), A Demand Theory for Number of Trips in a Random Utility Model of Recreation. *Journal of Environmental Economics and Management*, 29, 357-367.

Pattie, D.C. & J. Snyder (1996), Using a Neural Network to Forecast Visitor

Behaviour. *Annals of Tourism Research*, 23, 151-164.

Payne, J.W. (1976), Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis. *Organizational Behavior and Human Performance*, 16, 366-387.

Peterson, G.L., J.F. Dwyer & A.J. Darragh (1983), A Behavioral Urban Recreation Site Choice Model. *Leisure Sciences*, 6, 61-81.

Quinlan, J.R. (1993), *C4.5 Programs for Machine Learning*. San Mateo (CA): Morgan Kaufmann Publishers.

Raaijmakers, S.F.J.M. (1997), *Een Dag van Verschil*. Ph.D.-thesis, Tilburg University.

Schiffman, L.G. & L. Lazar Kanuk (1991), *Consumer Behavior*. Englewood Cliffs, New Yersey: Prentice-Hall.

Schmeink, H. & S.J. ten Wolde (1999), Recreatiegedrag en -Wensen van Alloctonen in Beeld. *Recreatie & Toerisme*, January 1999, 6-9.

Schmidhauser, H. (1976), Neue Erkenntnisse über Gesetzmässigkeiten bei der Wahl des Reiseziels. *Jahrbuch für Fremdenverkehr*, 24/25, 86-102.

Sewrajsing, A.S. (1996), *Buitenlandse Toeristen in Nederland.* Delft: TNO Infrastructuur, Transport en Regionale Ontwikkeling.

Shannon, R.E. (1975), *Systems Simulation. The Art and Science*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Shechter, M. & R.C. Lucas (1978), *Simulation of Recreational Use for Park and Wilderness Management*. Baltimore/London: The Johns Hopkins University Press for Resources for the Future.

Siderelis, C. & R.L. Moore (1998), Recreation Demand and the Influence of Site Preference Variables. *Journal of Leisure Research*, 30, 301-318.

Smith, S.L.J. (1995), *Tourism Analysis: A Handbook*. London: Longman.

Smith, V.K. & J.V. Krutilla (1976), *Structure and Properties of A Wilderness Travel Simulator: An Application to the Spanish Peaks Area*. Baltimore: Johns Hopkins University Press of Resources for the Future.

SPSS (1998), *AnswerTree 2.0 User's Guide*. Chicago, IL, USA.

Stemerding, M.P. (1996), *Modeling Constraints-Based Choices for Leisure Mobility Planning*. Ph.D.-thesis, Eindhoven University of Technology.

Stewart, R.J. & J. Matthews (1995), *Merlin Through the Ages. A Chronological Anthology and Source Book*. London: Bandford.

Stichting Toerisme & Recreatie AVN & Ministry of OC and W (1997), *Schoolvakanties 1997-2002*.

Strambi, O. & K.A. van de Bilt (1998), Trip Generation Modeling Using CHAID, a Criterion-Based Segmentation Modeling Tool. *Transportation Research*

Record 1645, 24-31.

Stynes, D.J. (1983), An Introduction to Recreation Forecasting. In: S.R. Lieber & D.R. Fesenmaier (Eds.), *Recreation Planning and Management* (pp. 87-95), London: E. & F.N. Spon Ltd.

Stynes, D.J. & G.J. Peterson (1986), A Review of Logit Models with Implications for Modeling Recreation Choices. *Journal of Leisure Research*, 16, 295-310.

Taha, I.A. & J. Ghosh (1999), Symbolic Interpretation of Artificial Neural Networks. IEEE Transactions on Knowledge and Data Engineering, 11(3, May/June), 448-463.

Te Kloeze, J.W. (1990), *Changing Economic Conditions in Relation to Recreation Styles in The Netherlands*. Paper presented to the “Leisure and the Future” Conference, Zajaczkowo, Poland.

Thill, J.C. & A. Wheeler (2000), Knowledge Discovery and Induction of Decision Trees in Spatial Decision. In: A. Regianni (Ed.), *Spatial Science: New Frontiers in Theory and Methodology*. Nuremberg: Springer.

Thurstone, L.L. (1927a), A Law of Comparative Judgement. *Psychological Review*, 34, 273-286.

Thurstone, L.L. (1927b), Psychological Analysis. *American Journal of Psychology*, 28, 368-389.

Timmermans, H.J.P. (1985a), Individuele Preferentiemodellen en Recreatiegedrag: Een Validatietoets. *Recreatie & Toerisme*, 447-451.

Timmermans, H.J.P. (1985b), Modellen van Recreatie-Gedrag: een Overzicht van Enkele Recente Ontwikkelingen. *Recreatie & Toerisme*, 390-394.

Timmermans, H.J.P. (2000), Theories and Models of Activity Patterns. In: T.A Arentze & H.J.P. Timmermans (Eds.), *ALBATROSS: A Learning Based Transportation Oriented Simulation System* (pp. 6-70), Eindhoven: EIRASS.

Timmermans, H.J.P. & A.W.J. Borgers (1989), Dynamic Models of Choice Behaviour: Some Fundamentals and Trends. In: J. Hauer *et al*. (Eds.), *Urban Dynamics and Spatial Choice behaviour* (pp. 3-26): Kluwer Academic Publishers.

Timmermans, H.J.P., A.W.J. Borgers, & P. van der Waerden (1991), Mother Logit Analysis of Substitution Effects in Consumer Shopping Destination Choice. *Journal of Business Research*, 23, 311-323.

Timmermans, H.J.P. & R.G. Golledge (1990), Application of Behavioural Research on Spatial Problems II: Preference and Choice. *Progress in Human Geography*, 14, 311-355.

Timmermans, H.J.P. & R.E.C.M. van der Heijden (1987), Uncovering Spatial Decision-Making Processes: A Decision Net Approach Applied to

Recreational Choice Behaviour. *Tijdschrift voor Economische en Socociale Geografie*, 78, 297-304.

Timmermans, H.J.P. & P. van der Waerden (1998), *Space-Time Consumption on the Move*. Eindhoven: EIRASS, in commission of the Projectbureau Integrale Verkeers- en Vervoerstudies (PbIvvs).

Train, K.E. (1998), Recreation Demand Models with Taste Differences Over People. *Land Economics*, 74, 230-39.

Trick, M.A. (1996), *An Introduction to Simulation*. Source: http://mat.gsia.cmu. edu.simul/node1.html

Um, S. & J.L. Crompton (1990), Attitude Determinants in Tourism Destination Choice. *Annals of Tourism Research*, 17, 432-448.

Urry, J. (1990), *The Tourist Gaze: Leisure and Travel in Contemporary Societies*. London: Sage Publications.

Uysal, M. & M.S. El Roubi (1999), Artificial Neural Networks versus Multiple Regression in Tourism Demand Analysis. *Journal of Travel Research*, 38, 111-118.

Van der Heijden, R.E.C.M. & H.J.P. Timmermans, (1987), Variatie in het Bezoek aan Natuur- en Recreatiegebieden door Openluchtrecreanten. *Vrijetijd en Samenleving*, 5, 115-140.

Van der Most, K. (1996), Continu Vakantie Onderzoek (CVO), *Handboek Recreatie en Toerisme* (pp. 1730-3 - 1730-12).

Van Dijk, J.C., J.S. Hagens, F.A.G. Windmeijer & J.W. Velthuijsen (1991), *A Simulation Model of the Dutch Tourist Market: A Technical Report on 'Toermodel'*. Amsterdam: Foundation for Economic Research of the University of Amsterdam (SEO) in commission by the Ministry of Economic Affairs, Department of Tourism.

Van Lier, H.N. (1993), New Challenges in Recreation and Tourism Planning. In: H.N. van Lier & P.D. Taylor (Eds.), *New Challenges in Recreation and Tourism Planning* (pp. 5-15), Amsterdam: Elsevier.

Van Keken, G.E. (1995), J. Lengkeek, B.H.M. Elands (1995), Nieuwe Dimensies in Zoneren: Zonering van Belevingswerelden. *Recreatie & Toerisme*, Dec., 6-9.

Van Middelkoop, M. (1996), *De relatie tussen vrijetijdsdoelen en openluchtrecreatief gedrag: restricted to the number of activities?* Interne Mededeling 418, DLO-Staring Centrum Wageningen.

Van Middelkoop, M., T.A. Arentze. A.W.J. Borgers & H.J.P. Timmermans (2000), Representing and Predicting Tourist Choice Behaviour: A Rule-Based vs. a Utility-Based Approach. *Tourism Analysis*, 5, 113-118.

Van Middelkoop, M., T.A. Arentze. A.W.J. Borgers & H.J.P. Timmermans (2001), Representing and Predicting Tourist Choice Behaviour: A Rule-Based vs. a Utility-Based Approach. In: J.A. Mazanec, G.I. Crouch, J.R. Brent Ritchie & A.G. Woodside (Eds.), *Consumer Psychology of Tourism, Hospitality and Leisure Volume Two* (pp. 107-121): CABI Publishing.

Van Middelkoop, M., A.W.J. Borgers & H.J.P. Timmermans (1999), Complementarity, Substitution and Independence among Tourist Trips. *Tourism Analysis*, 4, 63-74.

Vanthienen, J. (1994), Technical Correspondence: A More General Comparison of Decision Tables and Trees. *Communications of the ACM*, 37, 109-113.

Velthuijsen, J.W. & M. Verhagen (1994), A Simulation Model of the Dutch Tourist Market. *Annals of Tourism Research*, 21, 812-827.

Verhelst, N.P.M. (1993), *Beslissingstabellen*. Leidschendam: Lansen Publishing bv.

Wasser, B. (1998), Nederlanders Reizen niet Milieubewust. *Recreatie & Toerisme*, 34-36.

Wets, G. (1998), *Decision Tables in Knowledge-Based Systems: Adding Knowledge Discovery and Fuzzy Concepts to the Decision Table Formalism*. Ph.D.-thesis, Eindhoven University of Technology.

Wets, G., K, Vanhoof, T.A. Arentze & H.J.P. Timmermans (2000), Identifying Decision Structures Underlying Activity Patterns: An Exploration of Data Mining Algorithms. *Transportation Research Record*, 1718, 1-9.

Wijker, L.J. (1998), *Vakantiekeuzen van Nederlanders. Marketingpaper*. Breda/ Voorschoten: NRIT/Toerisme & Recreatie AVN.

Witlox, F.J.A. (1995), Qualitative Housing Choice Modelling: Decision Plan Nets Versus Decision Tables. *Netherlands Journal of Housing and the Built Environment*, 10, 209-237.

Witlox, F.J.A. (1998), *Modelling Site Selection*. Ph.D.-thesis, Eindhoven University of Technology.

Witt, C.A. & P.L. Wright (1993), Tourist Motivation: Life after Maslow. In: P. Johnson & B. Thomas (Eds.), *Choice and Demand in Tourism* (pp. 33-56), London: Mansell Publishing Limited.

Woodside, A.G. & S. Lysonski (1989), A General Model of Traveler Destination Choice. *Journal of Travel Research* 27, 4, 8-14.

Woodside, A.G. & R. MacDonald (1994), General System Framework of Customer Choice Processes of Tourism Services. In: R. Gasser & K. Weiermair (Eds.), *Spoilt for Choice* (pp. 31-59), Austria: Kultur Verlag.

Yamamoto, T. & R. Kitamura (1999), An Analysis of Time Allocation to In-Home and Out-of-Home Discretionary Activities Across the working Days and Non-Working Days. *Transportation*, 26, 211-230.

APPENDIX 1: Condition Variables

This appendix presents the condition variables that were entered in to the CHAID-based algorithm for the choice facets. The coding presented in this appendix corresponds to the coding used in the decision tables and is based on the original CVO-coding as much as possible.

A1.1 Introduction

This appendix introduces the key to the condition variables that are entered into the CHAID-based algorithm for each choice facet in the decision-making process. To find the set of conditions that is used for a particular facet, three tables should be combined. First, Table A1.1 comprises the conditions that are always included. Second, Table A1.2 presents the list of conditions that reflect previous decisions regarding the activity under consideration (based on the sequential decision making process with regard to trip profiling choices introduced in chapter 5). Conditions that are presented *above* the facet under consideration are entered. E.g., for the choice of travel party *none* of the conditions in Table A1.2 is used, whereas for expenditures *all* are entered. Third and finally, section A1.3 introduces, for each choice facet in turn, a table presenting the conditions that are specific to that choice facet.

Notes to all tables presented in this appendix:
(1) The column indicated by "T" records the type of condition variable: (N) Categorical/Nominal; (O) Ordinal; (C) Continuous; and (F) Floating
(2) o.m.i.t. means *of/on more important trips* (in the annual trip program)

A1.2 General conditions

Table A1.1 Conditions common to all choice facets

Label	T	Condition, States and Codes
		Characteristics of the annual Trip Program & the Trip
#DT	F	Annual No. of Day-Trips: (0) 0 day-trips; (1) 1-5 day-trips; (2) 6-10 day-trips; (3) 11-15 day-trips; (4) 16-20 day-trips; (5) 21-25 day-trips; (6) 26-30 day-trips; (7) 31-35 day-trips; (8) 36-40 day-trips; (9) 41-45 day-trips; (10) 46-50 day-trips; (11) 51 or more Day-Trips; (98) younger than 16 years; (99) unknown
#Hol	C	Annual No. of Holidays: 1-47 Holidays
Lngth	O	Duration of the trip: (1) short break: 2-4 days; (2) medium long holiday: 5-8 days; (3) extended holiday: 9-15 days; (4) long holiday: 16-28 days; (5) extra long holiday: 29+ days

Table A1.1 Continued - Conditions common to all choice facets

Label	T	Condition, States and Codes
		Personal and Household Characteristics
Age	C	0-98 years
Inc	O	Net Annual Household Income in 8 categories: (1) < € 10,210; (2) € 10,210-13,613; (3) € 13,613-15,882; (4) € 15,882-18,151; (5) € 18,151-20,420; (6) € 20,420-24,958; (7) € 24,958-29,496 (8) > € 29,496
Edu	O	Education in 9 levels: (0) no school yet .. to (8) highest
Gndr	N	Gender: (1) female; (2) male
Cst	N	Civil status: (1) Married; (2) Cohabiting; (3) Divorced; (4) Widow(er); (5) Single; (6) Other
Chld	N	Children in Household: (1) no children under 18; (2) (some) children aged 6-17; (3) children aged 0-5 only
HHsz	O	No. of people in household: (1) single ... to (10) 10 or more
SClss	O	Social Class in 5 levels: (1) highest ... to (5) lowest
Work	N	(0) younger than 16; (1) 30+ Hrs/week; (2) 1-30 Hrs/week; (3) no job
Wrkhr	C	No. of working hours per week: 0-99Hrs/week
WrkW	F	Work during weekend: (0) younger than 16; (1) never; (2) sometimes (1-5 x p.y.); (3) regularly (6-12 x); (4) often (13 +); (9) unknown
Freed	F	No. of free days in 1998; 0-98 days; (99) unknown, (100) younger than 16; (101) no job
Wrkrs	N	(1) respondent's job prescribes certain holiday seasons (e.g., teacher or nurse); (0) no restrictions
Region	N	School Holiday Region: (1) North; (2) Mid; (3) South
Prov	N	Province of residence: (1) Groningen; (2) Friesland; (3) Drenthe; (4) Overijssel; (5) Flevoland; (6) Gelderland; (7) Utrecht; (8) North-Holland; (9) South-Holland; (10) Zeeland; (11) North-Brabant; (12) Limburg
Urban	O	Urbanisation level city of residence: (1) very high... to (5) not urban
Csize	O	No. of inhabitants city of residence: (1.) less than 5.000; (2) 5-10.000; (3) 10-20.000; (4) 20-50.000; (5) 50-100.000; (6) 100-250.000; (7) 250.000 and more
Car	N	(1) car(s) household; (0) no car in household
Ski	N	(1) skis; (0) no skis
Fracc	N	(1) tourist accommodation (boat, tent, caravan) without a permanent place; (0) no accommodation without a permanent place
Peracc	N	(1) tourist accommodation (tent, caravan, 2nd house or boat) with a permanent place; (0) no accommodation with a permanent place

Table A1.2 Conditions reflecting previous decisions regarding the trip under consideration

Label	T	Condition, States and Codes
Party	N	Travel party: (1) alone; (2) adults only; (3) with schoolchildren (6-14 years); (4) with other children (0-5 or 15-19 years); (5) party of 9 or more people; (99) unknown
Prio4	O	Importance of the trip within the profiling process: (1) most important; (2) first but one; (3) first but two; (4) not in top-3
#Tp	C	No. of trips still to be profiled: 1-47 Holidays
#Dtot	C	No. of days o.m.i.t. already profiled
HolP	N	(1) in school holiday period; (0) outside school holiday period
Seas+	N	Season: (1) winter; (2) spring; (3) summer-begin; (4) summer-mid; (5) summer-end; (6) autumn
D_NLA	N	Main part of holiday: (1) in the Netherlands; (2) Abroad
D_NL	N	Domestic destination: (0) destination abroad; (1) Land north-Mid (Utrecht, 't Gooi and the Veluwe); (2) Water-North (Wadden Islands, the beaches of the Lake Yssel, lake areas in Groningen, Friesland and North-West Overijssel; (3) Water-Mid (Delta area, lake araes of Holland and Utrecht & river area of Gelderland); (4) Water-N-Sea (North Sea Coastal Area); (5) Land north-East (Achterhoek & Twente, Salland and the Vecht-area); (6) Land north-North (sand areas of Groningen, Friesland and Drenthe); (7) Land south-West (Western and Mid parts of Brabant); (8) Land south-South (Eastern part of Brabant, Northern and Mid parts of Limburg and realm of Nijmegen & southern part of Limburg); (9) Other (Amsterdam, Rotterdam, The Hague and Utrecht (cities) & remaining parts of the Netherlands)
D_A	N	Destination abroad: (0) domestic destination; (1) France; (2) Spain & Portugal; (3) Belgium & Luxembourg; (4) Germany; (5) Austria & Switzerland; (6) Italy & Greece; (7) England, Scotland, (Northern-) Ireland & Wales; (8) Former Yugoslavia & Soviet Union, Hungary, Czech Republic, Slovakia, Poland, Romania and Bulgaria; (9) Turkey, Morocco, Malta, Cyprus Tunisia & Other Europe; (10) Norway, Sweden, Denmark & Finland; (11) Other
Acc	N	Accommodation: (1) hotel, motel etc.; (2) non-permanent, owned by tourist; (3) permanent, owned by tourist; (4) other
Tmode	N	Means of transport: (1) car/camper; (2) aeroplane; (3) bus; (4) train/car(sleep)train & car; (5) other
DeptD	N	Departure date: (1) departure on Friday or Saturday; (0) other
Exp	C	Expenditures

A1.3 Conditions specific to each DT

Table A1.3 Conditions specific to the choice of “school holiday period or not”

Label	T	Condition, States and Codes
#Dih	C	No. of days o.m.i.t. already scheduled in school holiday period
#Dnh	C	No. of days o.m.i.t. already scheduled outside school holiday period
#Tih	C	No. o.m.i.t. already scheduled in school holiday period
#Tnh	C	No. o.m.i.t. already scheduled outside school holiday period

Table A1.4 Conditions specific to the choice of season

Label	T	Condition, States and Codes
#Dwnt	C	No. of days o.m.i.t. already scheduled in winter
#Dspr	C	No. of days o.m.i.t. already scheduled in spring
#Dsum	C	No. of days o.m.i.t. already scheduled in summer
#Daut	C	No. of days o.m.i.t. already scheduled in autumn
#Twnt	C	No. o.m.i.t. already scheduled in winter
#Tspr	C	No. o.m.i.t. already scheduled in spring
#Tsum	C	No. o.m.i.t. already scheduled in summer
#Taut	C	No. o.m.i.t. already scheduled in autumn

Table A1.5 Conditions specific to the choice of summer season period

Label	T	Condition, States and Codes
#Twnt	C	No. o.m.i.t. already scheduled in winter
#Tspr	C	No. o.m.i.t. already scheduled in spring
#Tsum	C	No. o.m.i.t. already scheduled in summer
#Taut	C	No. o.m.i.t. already scheduled in autumn
#Ts1	C	No. o.m.i.t. already scheduled in summer-begin
#Ts2	C	No. o.m.i.t. already scheduled in summer-mid
#Ts3	C	No. o.m.i.t. already scheduled in summer-end

Table A1.6 Conditions specific to the choice of departure date

Label	T	Condition, States and Codes
#Dfs	C	No. of days o.m.i.t. already leaving on Friday or Saturday
#Do	C	No. of days o.m.i.t. already leaving on Sun, Mon, Tue, Wed or Thu
#Tfs	C	No. o.m.i.t. already leaving on Friday or Saturday
#To	C	No. o.m.i.t. already leaving on Sun, Mon, Tue, Wed or Thu

Table A1.7 Conditions specific to all destination choices

Label	T	Condition, States and Codes
Pcar	C	Personal propensity to select car on a scale from 0 to 100
Pair	C	Personal propensity to select aeroplane on a scale from 0 to 100
Palt	C	Personal propensity to select alternative modes on a scale from 0 to 100
Tcost	C	Total amount of expenditure o.m.i.t.

Note: together the first three variables always sup up to 100

Table A1.8 Conditions specific to the choice of domestic or abroad

Label	T	Condition, States and Codes
#DA	C	No. of days o.m.i.t. already at foreign destinations
#DD	C	No. of days o.m.i.t. already at domestic destinations
#TA	C	No. o.m.i.t. already at foreign destinations
#TD	C	No. o.m.i.t. already at domestic destinations

Table A1.9 Conditions specific to the choice of neighbouring or more distant countries

Label	T	Condition, States and Codes
#AD	C	No. o.m.i.t. already at domestic destinations
#DNei	C	No. of days o.m.i.t. already at neighbouring foreign destinations
#DRem	C	No. of days o.m.i.t. already more distant foreign destinations
#Tnei	C	No. o.m.i.t. already at neighbouring foreign destinations
#TRem	C	No. o.m.i.t. already at more distant foreign destinations

Table A1.10 Conditions specific to the choice of the more distant destinations

Label	T	Condition, States and Codes
#TD	C	No. o.m.i.t. already at domestic destinations
#Tnei	C	No. o.m.i.t. already at neighbouring foreign destinations
#TRem	C	No. o.m.i.t. already at more distant foreign destinations
#TSpP	C	No. o.m.i.t. already in Spain/Portugal
#TSAu	C	No. o.m.i.t. already in Switzerland/Austria
#TItGr	C	No. o.m.i.t. already in Italy/Greece
#TFEB	C	No. o.m.i.t. already in Former Easter European Bloc countries
#TTurk	C	No.o.m.i.t. already in South Each Mediterranean Countries
#TScan	C	No. o.m.i.t. already in Norway/Sweden/Denmark/Finland
#TORD	C	No. o.m.i.t. already in other more distant destinations

Table A1.11 Conditions specific to the choice of the neighbouring countries

Label	T	Condition, States and Codes
#TD	C	No. o.m.i.t. already at domestic destinations
#Tnei	C	No. o.m.i.t. already at neighbouring foreign destinations
#TRem	C	No. o.m.i.t. already at more distant foreign destinations
#TFra	C	No. o.m.i.t. already in France
#TBLx	C	No. o.m.i.t. already in Belgium/Luxembourg
#TUK	C	No. o.m.i.t. already in UK/Ireland
#TGer	C	No. o.m.i.t. already in Germany

Table A1.12 Conditions specific to the choice of the Dutch metaregions

Label	T	Condition, States and Codes
#TD	C	No. o.m.i.t. already at domestic destinations
#Tnei	C	No. o.m.i.t. already at neighbouring foreign destinations
#TRem	C	No. o.m.i.t. already at more distant foreign destinations
#DWat	C	No. of days o.m.i.t. already in Dutch metaregion “Water”
#DLNo	C	No. of days o.m.i.t. already in Dutch metaregion “Land north”
#DLSo	C	No. of days o.m.i.t. already in Dutch metaregion “Land south”
#DOth	C	No. of days o.m.i.t. already inDutch metaregion “Other”
#TWat	C	No. o.m.i.t. already in Dutch metaregion “Water”
#TLNo	C	No. o.m.i.t. already in Dutch metaregion “Land north”
#TLSo	C	No. o.m.i.t. already in Dutch metaregion “Land south”
#TOth	C	No. o.m.i.t. already in Dutch metaregion “Other”

Table A1.13 Conditions specific to the choice of the tourist areas within the Dutch metaregion “Land South”

Label	T	Condition, States and Codes
#TD	C	No. o.m.i.t. already at domestic destinations
#Tnei	C	No. o.m.i.t. already at neighbouring foreign destinations
#TRem	C	No. o.m.i.t. already at more distant foreign destinations
#TWat	C	No. o.m.i.t. already in Dutch metaregion “Water”
#TLNo	C	No. o.m.i.t. already in Dutch metaregion “Land north”
#TLSo	C	No. o.m.i.t. already in Dutch metaregion “Land south”
#TLSo	C	No. o.m.i.t. already in Dutch metaregion “Other”
#TLSW	C	No. o.m.i.t. already in Dutch region “Land south-West”
#TLSS	C	No. o.m.i.t. already in Dutch region “Land south-South”

Table A1.14 Conditions specific to the choice of the tourist areas within the Dutch metaregion “Land North”

Label	T	Condition, States and Codes
#TD	C	No. o.m.i.t. already at domestic destinations
#Tnei	C	No. o.m.i.t. already at neighbouring foreign destinations
#TRem	C	No. o.m.i.t. already at more distant foreign destinations
#TWat	C	No. o.m.i.t. already in Dutch metaregion “Water”
#TLNo	C	No. o.m.i.t. already in Dutch metaregion “Land north”
#TLSo	C	No. o.m.i.t. already in Dutch metaregion “Land south”
#TLSo	C	No. o.m.i.t. already in Dutch metaregion “Other”
#TLNN	C	No. o.m.i.t. already in Dutch region “Land north North”
#TLNM	C	No. o.m.i.t. already in Dutch region “Land north-Mid”
#TLNE	C	No. o.m.i.t. already in Dutch region “Land north -East”

Table A1.15 Conditions specific to the choice of the tourist areas within the Dutch metaregion “Water”

Label	T	Condition, States and Codes
#TD	C	No. o.m.i.t. already at domestic destinations
#Tnei	C	No. o.m.i.t. already at neighbouring foreign destinations
#TRem	C	No. o.m.i.t. already at more distant foreign destinations
#TWat	C	No. o.m.i.t. already in Dutch metaregion “Water”
#TLNo	C	No. o.m.i.t. already in Dutch metaregion “Land north”
#TLSo	C	No. o.m.i.t. already in Dutch metaregion “Land south”
#TLSo	C	No. o.m.i.t. already in Dutch metaregion “Other”
#TWN	C	No. o.m.i.t. already in Dutch region “Water-North”
#TWM	C	No. o.m.i.t. already in Dutch region “Water-Mid”
#TWNs	C	No. o.m.i.t. already in Dutch region “Water-North Sea Coastal Area”

Table A1.16 Conditions specific to the choice of accommodation

Label	T	Condition, States and Codes
Tcost	C	Total amount of expenditure o.m.i.t.
#DHot	C	No. of days o.m.i.t. already in hotel/motel
#DNper	C	No. of days o.m.i.t. already in personal non-permanent acc.
#Dper	C	No. of days o.m.i.t. already in personal permanent acc.
#DOA	C	No. of days o.m.i.t. already in other accommodations
#THot	C	No. o.m.i.t. already in hotel/motel
#TNper	C	No. o.m.i.t. already in personal non- permanent acc.
#Tper	C	No. o.m.i.t. already in personal permanent acc.
#TOA	C	No. o.m.i.t. already in other accommodations

Table A1.17 Conditions specific to the choice of transport mode

Label	T	Condition, States and Codes
Tcost	C	Total amount of expenditure o.m.i.t.
Pcar	C	Personal propensity to select car on a scale from 0 to 100
Pair	C	Personal propensity to select aeroplane (0-100)
Palt	C	Personal propensity to select alternative modes (0-100)
#Dcar	C	No. of days o.m.i.t. already by car
#Dair	C	No. of days o.m.i.t. already by aeroplane
#Doth	C	No. of days o.m.i.t. already by other modes of transport
#Tcar	C	No. o.m.i.t. already by car
#Tair	C	No. o.m.i.t. already by aeroplane
#Toth	C	No. o.m.i.t. already by other modes of transport

Table A1.18 Conditions specific to the choice of “alternative” transport modes

Label	T	Condition, States and Codes
#Dtrain	C	No. of days o.m.i.t. already by train
#Dbus	C	No. of days o.m.i.t. already by coach/shuttle bus
#DOoth	C	No. of days o.m.i.t. already by other-other modes of transport
#Ttrain	C	No. o.m.i.t. already by train
#Tbus	C	No. o.m.i.t. already by bus
#TOoth	C	No. o.m.i.t. already by other-other modes of transport

Table A1.19 Conditions specific to the choice of expenditures

Label	T	Condition, States and Codes
Tcost	C	Total amount of expenditure o.m.i.t.

APPENDIX 2: Sensitivity Analyses

This appendix presents the sensitivity analyses that were performed to determine the sensitivity to stopping criteria of the CHAID-based rule-induction algorithm for the choice the various tourist choice facets. All analyses were conducted using an all-or-nothing assignment of observed choices and a 4-fold cross-over validation technique. The stopping criteria that were selected for the induction of tourists' decision rules are marked in bold.

A2.1 Travel party choices

Table A2.1 The choice of travel party (all 7121 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	134	70.24%	70.50%	-- [a]
50/20	116	71.63%	70.51%	-- [a]
60/25	107	71.24%	70.43%	-- [a]
80/35	74	70.51%	70.05%	--- [ab]
100/45	**68**	**70.20%**	**69.72%**	**-- [a]**

[a] the model seriously underestimates the number of independent travellers and the number of travel parties of at least nine people; it overestimates the number of travel parties with adults only

[b] none of the rules is able to identify travel parties of at least nine people

A2.2 Tourist timing choices

Table A2.2 The choice of "school holiday period or not" (all 7121 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	65	67.06%	62.73%	- [a]
50/20	58	67.18%	62.79%	+/- [a]
60/25	50	66.72%	62.69%	+/- [a]
80/35	**46**	**66.67%**	**62.94%**	**+/- [a]**
100/45	38	66.27%	62.93%	-- [a]

[a] the model (slightly) underestimates the number of trips outside school holiday periods

Table A2.3 The choice of season (all 7121 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	121	65.58%	62.24%	-- [a]
50/20	98	64.79%	62.37%	-- [a]
60/25	**88**	**65.50%**	**62.44%**	**-- [a]**
80/35	68	64.29%	62.15%	--- [a]
100/45	60	63.88%	62.34%	--- [a]

[a] the model seriously underestimates the number of autumn and winter trips, and overestimates the more popular spring and summer seasons; this bias increases with stricter stopping criteria

Table A2.4 The choice of summer season period (2612 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	36	54.33%	48.97%	+ [a]
50/20	30	53.64%	48.89%	+/- [a]
60/25	29	53.41%	48.77%	+/- [a]
80/35	**24**	**53.33%**	**49.58%**	**+/- [a]**
100/45	23	53.25%	49.46%	+/- [a]

[a] the model (slightly) overestimates the number of trips during the mid- and end-summer period

Table A2.5 The choice of departure date (all 7121 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	113	71.25%	64.50%	- [b]
50/20	99	70.72%	65.24%	- [b]
60/25	91	70.14%	65.00%	- [b]
80/35	76	69.99%	65.31%	+ [a]
100/45	**68**	**69.43%**	**65.80%**	**+ [a]**

[a] the model slightly overestimates departure on Fridays and Saturdays

[b] the model underestimates departure on Fridays and Saturdays

A2.3 Tourist destination choices

Table A2.6 The choice of domestic or abroad (all 7121 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	116	78.25%	73.54%	-[a]
50/20	102	77.57%	73.91%	+/-[a]
60/25	89	77.00%	74.10%	+/-[a]
80/35	**74**	**76.90%**	**73.70%**	**+/-[a]**
100/45	60	76.17%	73.54%	-[a]

[a] the model slightly overestimates the number of domestic trips

Table A2.7 The choice of neighbouring or more distant country (3266 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	50	73.36%	69.14%	+[a]
50/20	46	73.52%	69.17%	+/-[a]
60/25	40	73.33%	68.71%	+[a]
80/35	**34**	**72.35%**	**68.71%**	**++**
100/45	27	71.43%	67.52%	+[b]

[a] the model slightly overestimates the number of “other abroad” trips

[b] the model slightly overestimates the number of “neighbouring country” trips

Table A2.8 The choice of the 7 more distant tourist destinations (1610 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	**33**	**42.48%**	**34.91%**	**--[a]**
50/20	29	41.80%	34.91%	---[ab]
60/25	27	41.24%	35.53%	---[ab]

[a] the model overestimates the number of trips to Spain/Portugal and Austria/ Switserland and seriously underestimates the number of trips to Scandinavia/ Denmark, South East Mediterranean countries and “other countries”; the model is unable to identify trips to the former East Bloc

[b] as ([a]), the model is unable to identify trips to Scandinavia/Denmark

Table A2.9 The choice of the 4 neighbouring countries (1656 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	**40**	**54.81%**	**45.95%**	**-[a]**
50/20	31	54.47%	48.43%	---[a]
60/25	30	54.23%	48.31%	---[a]

[a] at the aggregate level, the model overestimates the number of trips to France, Belgium & Luxembourg and Germany and seriously underestimates the number of trips to the UK/Ireland

Table A2.10 The choice of the 4 Dutch metaregions (3855 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	76	62.78%	54.45%	--[a]
50/20	65	62.05%	54.24%	---[a]
60/25	56	61.58%	54.24%	---[a]
80/35	44	61.01%	54.01%	--[a]
100/45	**34**	**58.99%**	**55.02%**	**--[a]**

[a] the model overestimates the number of trips to the regions “Water” and “Land north”, and (seriously) underestimates the number of trips to the regions “Land south” and “Other”

Table A2.11 The choice of the tourist areas within the Dutch metaregion “Land south” (976 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	14	78.59%	74.18%	-[a]
50/20	**14**	**78.79%**	**75.20%**	**-[a]**
60/25	12	77.05%	74.90%	-[a]

[a] the model slightly overestimates the trips to “South” and underestimated the trips to “West”

Table A2.12 The choice of the tourist areas within the Dutch metaregion "Land north" (1491 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	21	59.95%	55.60%	--[a]
50/20	16	59.15%	55.67%	-[a]
60/25	**14**	**58.35%**	**55.80%**	**-[a]**

[a] the model overestimates the trips to "Mid" and (seriously) underestimated the trips to "East" and "North"

Table A2.13 The choice of the tourist areas within the Dutch metaregion "Water" (1177 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	19	67.54%	62.79%	-[a]
50/20	16	65.76%	62.62%	-[b]
60/25	**15**	**65.76%**	**63.47%**	**-[b]**

[a] the model overestimates the trips to "N-Sea" and "North" and seriously underestimated the trips to "Mid"

[b] the model overestimates the trips to "N-Sea" and seriously underestimated the trips to "Mid"

A2.4 Tourist accommodation choices

Table A2.14 The choice of accommodation (all 7121 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	151	76.24%	71.59%	+/-[a]
50/20	124	75.69%	71.58%	+/-[ab]
60/25	108	75.00%	71.59%	+/-[ab]
80/35	**90**	**74.12%**	**71.23%**	**+/-[a]**
100/45	78	73.78%	70.50%	-[a]

[a] the model slightly overestimates the number of "hotel, etc" choices and underestimates the number of "other" choices

[b] as (a), the model also overestimates the number of "not-fixed - privately owned" choices

A2.5 Tourist travel mode choices

Table A2.15 The choice of transport mode (all 7121 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	151	86.18%	84.99%	-[a]
50/20	124	86.59%	84.79%	-[ab]
60/25	**100**	**86.31%**	**85.06%**	**-[ab]**
80/35	88	86.11%	84.47%	--[ab]
100/45	72	85.37%	84.57%	--[a]

[a] the model overestimates the number of car-trips and underestimates the number of "alternative" transport modes
[b] as (a), the model also overestimates the number of aeroplane-based trips

Table A2.16 The choice of "alternative" transport modes (918 observations)

Before/After	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
35/15	**32**	**73.20%**	**66.56%**	**-[a]**
50/20	22	70.48%	65.89%	--[a]
60/25	20	69.83%	64.60%	-[a]

[a] the model overestimates the number of bus trips and underestimates the number of trips by train

A2.6 Tourist expenditure choices

Table A2.17 The choice of the segment of expenditures (all 7121 observations)

Before/After[a]	# rules	% correct		Aggregate performance
		Resubstitution	Cross-validation	
80/35	117	38.91%	33.16%	+/-[bcf]
100/45	98	37.92%	33.03%	+/-[cefg]
120/55	**79**	**36.60%**	**33.01%**	**+/-[cdg]**

[a] using stricter stopping criteria, the model becomes unmanageable
[b] the model overestimates P2 (NLG 55-120)
[c] the model overestimates P9 (NLG 1100-1750)
[d] the model underestimates P4 (> NLG 187.50-266.53)
[e] the model underestimates P6 (> NLG 375-500)
[f] the model underestimates P8 (NLG 750-1100)
[g] the model underestimates P10 (> NLG 1750)

APPENDIX 3: Weighting Schemes

This appendix presents the weighting schemes that are introduced to the sequential utility-based trip generation/time allocation models. Since two sets of sequential models were calibrated (see chapter 7), weighting schemes are induced for "adult" and the "child" models separately.

Notes to all tables:

98CVO: Observed participation and/or trip choices in the 1998 CVO data (for that segment of the sample with N_s respondents);

Uar (N_u) : Usable allocation responses (number of usable respondents N_u), representing the participation and/or trip choice responses to the additional questions that were used to calibrate the time allocation/ trip generation models;

Ψ_t : Weight for trip type *t* (*t* = "day-trip", "short break", "extended holiday", "long holiday" or "extra long holiday"); this weight bridges the differences between the "usable allocation responses" and the 1998 CVO-data (see chapter 6); based on the average number of trips p.p.p.y. Ψ_t is calculated as: Ψ_t = 98CVO / Uar;

$¥_t$: Additional weight for trip type *t* that ensures that the predicted trip frequencies match the observed number of trips (98CVO); based on the overall trip frequencies, $¥_t$ is calculated as: $¥_t$ = 98CVO / Predicted after Ψ_t

Y_t : Overall weight for trip type *t*: $Y_t = \Psi_t * ¥_t$

$¥_p$: Weight for the participation choice alternative *p* (*p* = 'day-trips only', 'holidays only' or ' both day-trips and holidays') that ensures that the predicted participation choices match the observed choices (both 98CVO and Uar) as closely as possible;

A3.1 "Adult" models

Table A3.1 Observed and predicted participation choices for adults in possession of a tourist accommodation with a permanent location (N_s = 215)

Participation choice alternative *p*	No tourist trips	Day-trips only	Holidays only	Both Day-trips and Holidays
98CVO	?	?	202 (94.0%)	
Uar (N_u = 97)	4 (4.12%)	3 (3.09%)	20 (20.6%)	70 (72.2%)
			90 (92.8%)	
Predicted[a]	14 (6.51%)	9 (4.19%)	52 (24.2%)	140(65.1%)
			192 (89.3%)	
Adjustment $¥_p$	-	.740	.880	1.19
Predicted after $¥_p$[a]	6 (2.80%)	7 (3.26%)	46 (21.4%)	156 (72.6%)
			202 (94.0%)	

[a] Average based on 1000 simulation runs

Table A3.2 Observed and predicted trip frequencies for adults in possession of a tourist accommodation with a permanent location (N_s =215) (average number of trips per person per year in brackets)

Trip type *t*	DT	SB	MLH	EH	LH	ELH	Total[b]
98CVO	?	706 (3.28)	264 (1.23)	145 (.674)	89 (.414)	54 (.251)	1258 (5.85)
Uar. (N_u = 97)	1606 (16.6)	337 (3.47)	102 (1.05)	58 (.598)	37 (.381)	10 (.103)	544 (5.61)
Weight Ψ_t	-	.945	1.17	1.13	1.09	2.44	-
Predicted after Ψ_t[ac]	3505 (16.3)	693	263	153	78	33	1135
Weight $¥_t$	1.02	1.02	1.00	.948	1.14	1.64	-
$Y_t = \Psi_t * ¥_t$	1.02	.963	1.17	1.07	1.24	3.99	
Predicted after Ψ_t and $¥_t$[ac]	3480 (16.2)	705	263	145	89	55	1256

[a] Average based on 1000 simulation runs
[b] Total number of holidays (excluding day-trips)
[c] Including the weight in the participation model ($¥_p$)

Table A3.3 Observed and predicted participation choices for adults in possession of a tourist accommodation without a permanent location (N_s = 915)

Participation choice alternative *j*	No tourist trips	Day-trips only	Holidays only	Both Day-trips and Holidays
98CVO	?	?	841 (91.9%)	
Uar. (N_u = 542)	16	27	91 (16.8%)	408 (75.3%)
	(2.95%)	(4.98%)	499 (92.1%)	
Predicted[a]	27	52	150 (16.4%)	6856(75.0%)
	(2.95%)	(5.68%)	836 (91.4%)	
Adjustment ¥$_p$	-	.930	1.01	1.00
Predicted after ¥$_p$[a]	29	49	151 (16.5%)	686 (75.0%)
	(3.17%)	(5.36%)	837 (91.5%)	

[a] Average based on 1000 simulation runs

Table A3.4 Observed and predicted trip frequencies for adults in possession of a tourist accommodation without a permanent location (N_s = 915; average number of trips per person per year in brackets)

Trip type *t*	DT	SB	MLH	EH	LH	ELH	Total[b]
98CVO	?	688 (.752)	596 (.651)	427 (.467)	396 (.433)	52 (.057)	2159 (2.36)
Uar (N_u = 542)	10456 (19.2)	530 (.978)	338 (.624)	233 (.430)	200 (.369)	27 (.050)	1328 (2.45)
Weight Ψ_t	-	.769	1.05	1.09	1.18	1.15	-
Predicted after Ψ_t[ac]	19361 (21.2)	781	579	425	392	39	2217
Weight ¥$_t$	.906	.881	1.03	1.00	1.01	1.33	-
$Y_t = \Psi_t$ * ¥$_t$	.906	.677	1.08	1.09	1.19	1.53	-
Predicted after Ψ_t and ¥$_t$[ac]	17529 (19.2)	687	597	428	395	52	2159

[a] Average based on 1000 simulation runs
[b] Total number of holidays (excluding day-trips)
[c] Including the weight in the participation model (¥$_p$)

Table A3.5 Observed and predicted participation choices for adults without a tourist accommodation (N_s =1706)

Participation choice alternative *p*	No tourist trips	Day-trips only	Holidays only	Both Day-trips and Holidays
98CVO	?	?	1196 (70.1%)	
Uar. (N_u =891)	175 (19.6%)	103 (11.6%)	148 (16.6%)	465 (52.2%)
			613 (68.8%)	
Predicted[a]	367 (21.5%)	208 (12.2%)	294 (17.2%)	837 (49.1%)
			1131 (66.3%)	
Adjustment $¥_p$	-	.950	.950	1.08
Predicted after $¥_p$[a]	326 (19.1%)	198 (11.6%)	278 (16.3%)	904 (53.0%)
			1182 (69.3%)	

[a] Average based on 1000 simulation runs

Table A3.6 Observed and predicted trip frequencies for adults without a tourist accommodation (average number of trips per person per year in brackets; N_s =1706)

Trip type *t*	DT	SB	MLH	EH	LH	ELH	Total[b]
98CVO	?	669 (.392)	801 (.470)	599 (.351)	272 (.159)	50 (.029)	2391 (1.40)
Uar (N_u = 891)	10803 (12.1)	529 (.594)	389 (.437)	315 (.354)	126 (.141)	11 (.012)	1370 (1.54)
Weight Ψ_t	-	.660	1.08	.992	1.13	2.42	-
Predicted after Ψ_t[ac]	22276 (13.1)	814	774	589	271	65	2513
Weight $¥_t$	.924	.822	1.03	1.02	1.00	.769	-
$Y_t = \Psi_t * ¥_t$	.924	.542	1.12	1.01	1.13	1.86	-
Predicted after Ψ_t and $¥_t$[ac]	20587 (12.1)	669	803	599	272	50	2393

[a] Average based on 1000 simulation runs
[b] Total number of holidays (excluding day-trips)
[c] Including the weight in the participation model ($¥_p$)

A3.2 “Child” models

Table A3.7 Observed and predicted participation choices for children (N_s=726)

Participation choice alternative *j*	No tourist trips or day-trips only	Holidays (with(out) day-trips)
Obs. (‘98 CVO)	114	612
Predicted[a]	114	612

[a] Average based on 1000 simulation runs

Table A3.8 Observed and predicted trip frequencies for children (N_s=726)

Type *t*	SB	MLH	EH	LH	ELH	Total
Obs. (‘98 CVO)	459	366	267	204	17	1313
Predicted[a]	459	367	267	203	17	1313

[a] Average based on 1000 simulation runs

Based on these simulation results it was decided not to introduce adjustment factors to the “child” models.

APPENDIX 4: *MERLIN* Manual

This appendix provides a guide to the *MERLIN*-system. Step-by-step this appendix takes the user through the various stages that are required to use the system sensibly - from the installation through to the interpretation of the output.

A4.1 Installation and default settings

To install the *MERLIN*-system on your computer, please insert the installation CD-rom. On this installation disc you will find a directory that contains the set-up program ('\Installing Merlin*.*'), a directory that contains two large databases that allow you to create cross tables based on the default reference project *MERLIN*-null ('\Merlin-null databases*.dbf'), a directory that contains the instructions to install Microsoft Agents on your computer ('\Installing MS agents*.pdf) and a directory that contains this thesis in pdf-format ('\Thesis*.pdf'). The *MERLIN*-system will be installed on your computer, simply by activating the set-up program ('\Installing Merlin\Setup.exe') and following the instructions presented to you. By default, the set-up program will install the *MERLIN*-system in the 'Program Files'-directory ('\Program Files\ Merlin\'); however, you are free to choose any other location as long as the files and subfolders within the *program directory* named '\Merlin' stay together and are not changed in any way (e.g., renamed, moved, and so on).

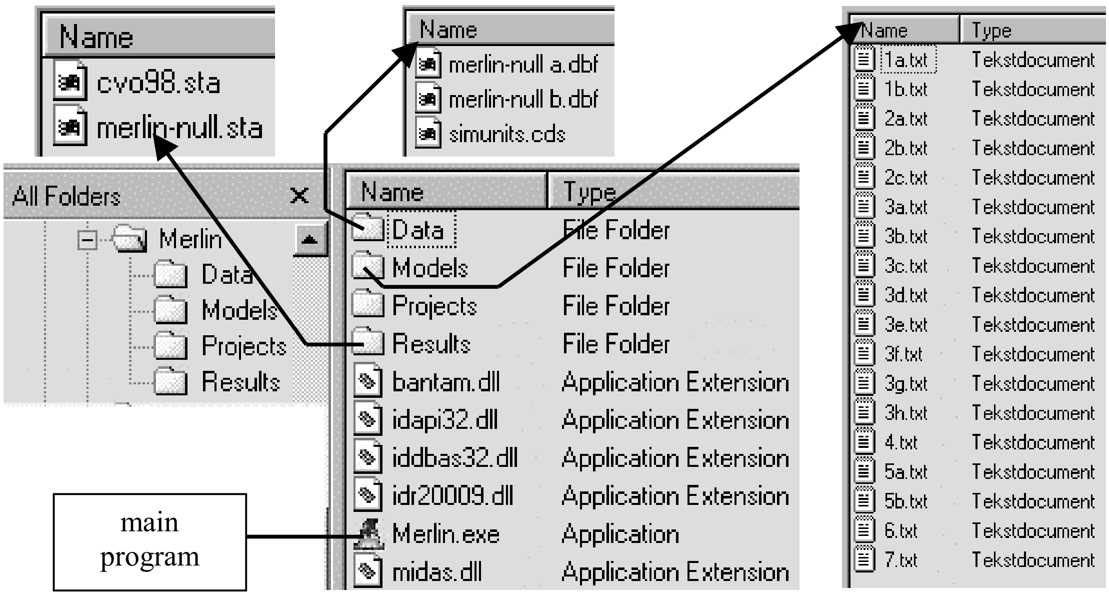


Figure A4.1 Directory structure and contents of the program, data, models and results directories, following the successful installation of the *MERLIN*-system

WARNING: Database-, project- and result-files should never be renamed or removed from the directory in which they are stored! Instead, when needed, define a new project based on the same data but a different name.

The resulting structure of the *program directory* named '\Merlin' is presented in Figure A4.1. This figure also includes two databases in the *data directory* that are not installed by default (see hereafter). In addition, the *.dll files in the program directory may differ depending on the settings of your computer.

Once you have successfully completed the installation process you can start working with the main program (see hereafter). However, there are two additional features that may increase the use and/or pleasure of working with the MERLIN-system. First, it is possible to further explore the results of the default reference project MERLIN-null (see chapter 12) by creating cross tables based on the simulated (N = 200 runs) tourist trip patterns of the 3562 members of 1998 CVO-panel. (This option is not incorporated in the set-up program by default because this requires two very large data bases - approximately 76 MB and 85 MB). If you want to use this option (either permanently or temporarily), please copy the two files in the database directory on your installation CD-rom ('\Merlin-null databases\Merlin-null a.dbf' and '\Merlin-null databases\Merlin-null b.dbf') into the *data directory* of the Merlin program directory ('\Program Files\Merlin\Data).

Second, you can meet and enjoy an agent named MERLIN while using the MERLIN-system. This will definitely make your day! If you want to meet MERLIN in person while using the system, you should obtain several Microsoft Agent Downloads. The instructions for installing the Microsoft Agent can be found on your installation CD-rom ('\Installing MS agents\Installing MS agents.pdf).

Once the system has been installed, Merlin.exe can be activated to define and simulate projects and study the results of these experiments. Figure A4.2 shows the first dialogue box that allows you to create a new project, open an existing project or, once you have finished your experiments, exit the system. If you want to open an existing project, the next dialogue box will show you the projects that are in the projects directory (*.pro files). Only one project can be activated at the time. Alternatively, if you want to create a new project, you will be asked to enter a name for this project. The default name is "New Project".

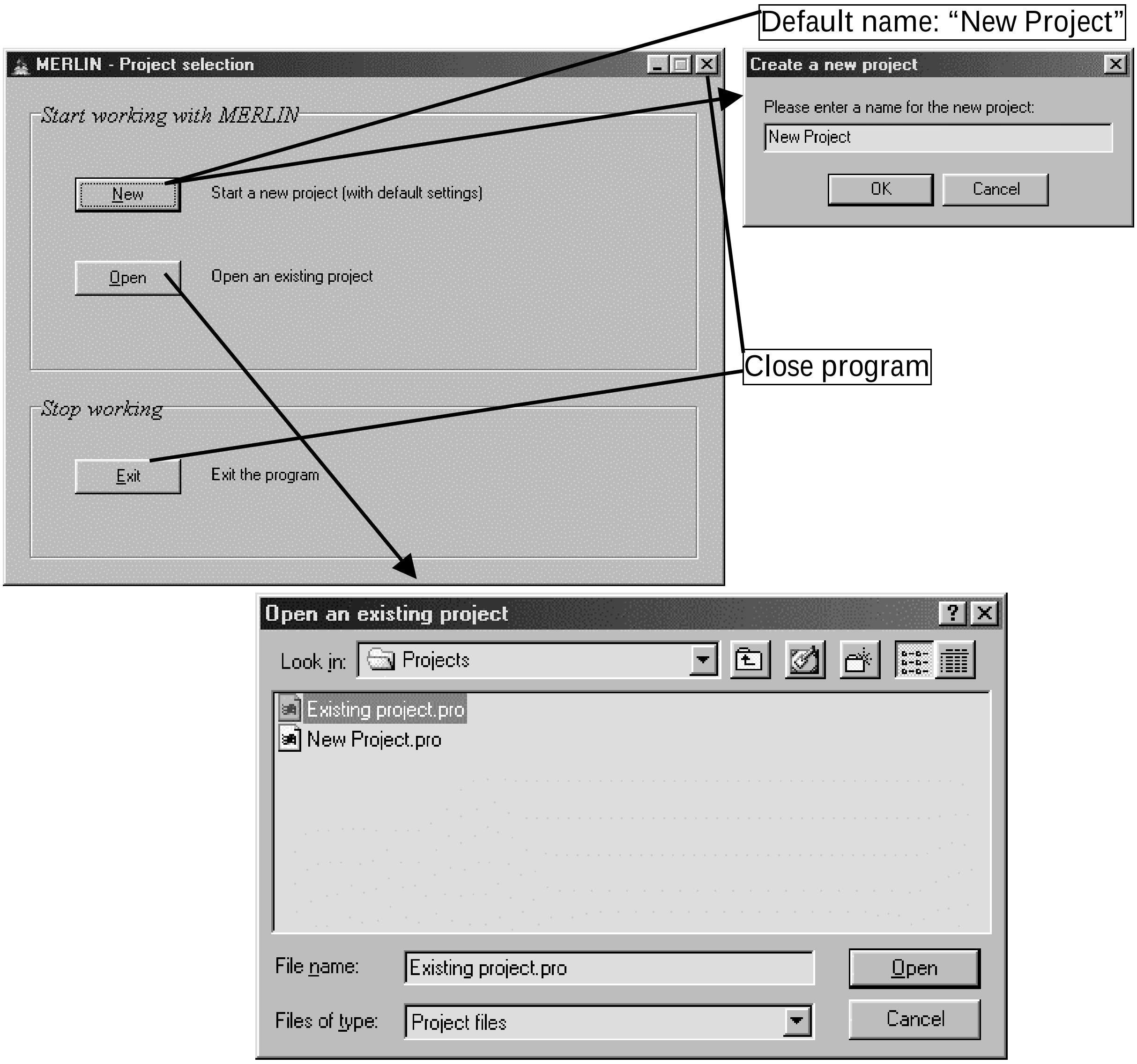


Figure A4.2 Dialogue boxes for opening new or existing projects and shutting down the *MERLIN*-system

Once you have provided a name for the (new) project, the next dialogue box will show the operations that can be performed, including: defining the project settings, defining the policy settings (i.e. the scenario for the future), run a simulation experiment (given the current settings), create cross-tables, and close the project. Figure A4.3 shows the dialogue box for these options. If a simulation experiment is run without further defining the project, the *default settings* will be used, i.e. the population of simulation units consists of the 3562 members of the 1998 CVO-panel (without any form of uprating or ageing), 10 simulations are run, the results of which are stored in data-bases (in addition to the standard output) and the reference project is "*MERLIN*-null". In this manual, the development of a simulation experiment is explained on the basis of a project named "Example".

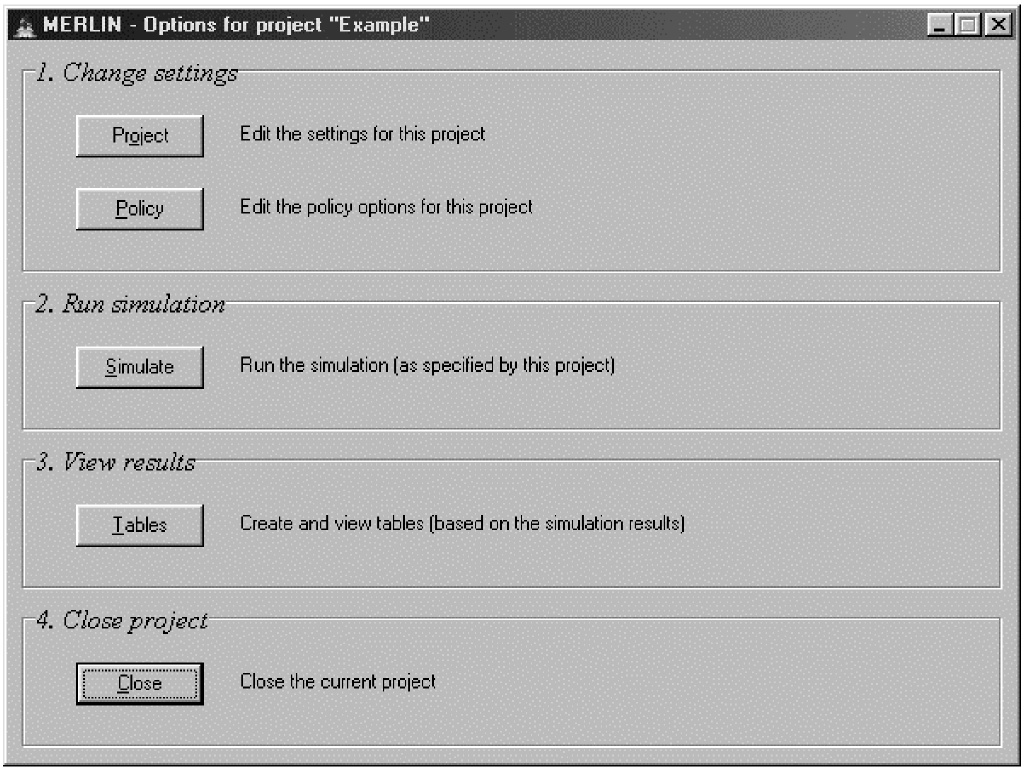


Figure A4.3 Dialogue box for project options

A4.2 Defining a project

Defining a project consists of two steps. First, the project settings should be defined. Second, the policy settings, i.e. the scenario for the future, should be specified.

A4.2.1 Project settings

The project settings include the name, the description, the number of simulation runs, the reference project, the possible division of the population of simulation units, and the (de)activation of the production of databases during the simulation process. Project settings can be changed by activating the “Project”-button in the dialogue box for project options (see Figure A4.3).

The *general project settings* include the name and the description of a project. By default, the name that was specified when the project was created is shown in the general project settings. This name will be used to identify the files that are related to this project. The extension marks the type and the contents of file:

- *.pro project file that stores the settings
- *.sta project file that stores the statistics of a project (it is used by the system

for creating the .res file; it is of no use to the user)

- *.res standard output file comprising several indicators of the simulation results at the aggregate level, including a (statistical) comparison to the reference project
- *.tab text-file containing a cross-table
- *.dsc project file that stores information with regard to the databases; it is of no use to the user;
- *a.dbf database file comprising the simulation results for each simulation unit for each run
- *b.dbf database file comprising the characteristics of each holiday that was generated and profiled during the simulation runs

If an existing file is opened, and it is not renamed, the existing database-, project- and result-files will be overwritten. This also applies to the default name "New Project" for new projects. The "description" part of the general project settings allows you to keep track of your work. These notes are included in the standard output file (*.res) of your project.

The *simulation settings* allow you to specify the number of simulation runs *N*. The default number is 10, but you are allowed to specify any number up to 200 runs. In addition, the project settings allow you to split the population of simulation units into two parts by specifying the proportions of the population (in %) that should be in each part. In this case, the population will be split up randomly into two parts A and B, each with the desired proportion of the population. In the standard output file (.res), the simulation results of part A will serve as the reference for the results of the units that are members of part B. This option allows you to assess the homogeneity of the population[44].

In addition to the general and simulation settings, the dialogue box for project settings also comprises output settings (results). This tab-leave allows you to select the reference project. In the standard output file (*.res) of the project under consideration, several indicators of the simulation results at the aggregate level are compared to this reference project. Indicators include, for instance, the average number of holidays over the *N* simulation runs, the average expenditures on domestic holidays over the *N* simulation runs, etc. In the *.res file, differences between the project under consideration and the reference project are tested using

[44] In chapter 12 it has been noted that in an experiment in which the population was randomly split into two equal halves, the simulation results of the two parts proved to differ significantly. It was concluded that differences should probably be attributed to the differences in the compositions of the two halves.

a *t*-test. The default reference project is "MERCIN-null" and represents the simulated (*N* = 200) tourist trip patterns of the 3562 members of 1998 CVO-panel without any form of upgrading and/or ageing.

Finally, the results settings allow you to deactivate the production of the databases (*.dbf files) during the execution of the simulation runs. These databases are required when you want to create cross-tables of two indicators. It is strongly recommended to deactivate the production of these databases when you do not intent to use the "Tables"-options because these databases tend to get very large even at very low numbers of simulation runs. Alternatively, you can delete the *.dbf files once you have finished your cross-tabulations. This will not affect any of the other results of the project under consideration.

A4.2.2 Policy settings

Once the project settings have been defined, you can specify the policy settings, i.e. the scenario for the future. By clicking the "Policy"-button in the dialogue box for project options (see Figure A4.3), the policy wizard that will guide you through the various options is activated. The first dialogue box (see Figure A4.4) allows you to select a particular segment of the population. By default, all respondents of the 1998 CVO-population are included in the population of simulation units. By ticking "Restrict the population to a selected segment", you can define segments of the population by either selecting a set (ANY) or a cross-section (ALL) of the population characteristics in the box on the right. In figure A4.4, for instance, a population is defined that consists of people from either single or two-person households. If you deselect the "Restrict the population to a selected segment"-function while a number of characteristics have already been selected, the population will not be restricted to that segment, but you can easily reactivate the selection because the information will be stored with your project. If you click "Cancel", however, the policy wizard will be closed and you will return to the options box for this project without storing your adjustments. If you want to delete the selection all together, click "Clear all".

When the preferred selection has been defined[45], the next dialogue box ("Next >") allows you to activate one of the reallocation scenarios that is presented in Figure A4.5 (see also chapters 6 and 7, and section 12.4.1).

[45] WARNING: If you want to assess the impact of a certain scenario for the future for a particular segment of the population, it is strongly recommended to first run a project without any changes for that segment. Next, this project can be used as the reference for the project that includes the changes under consideration.

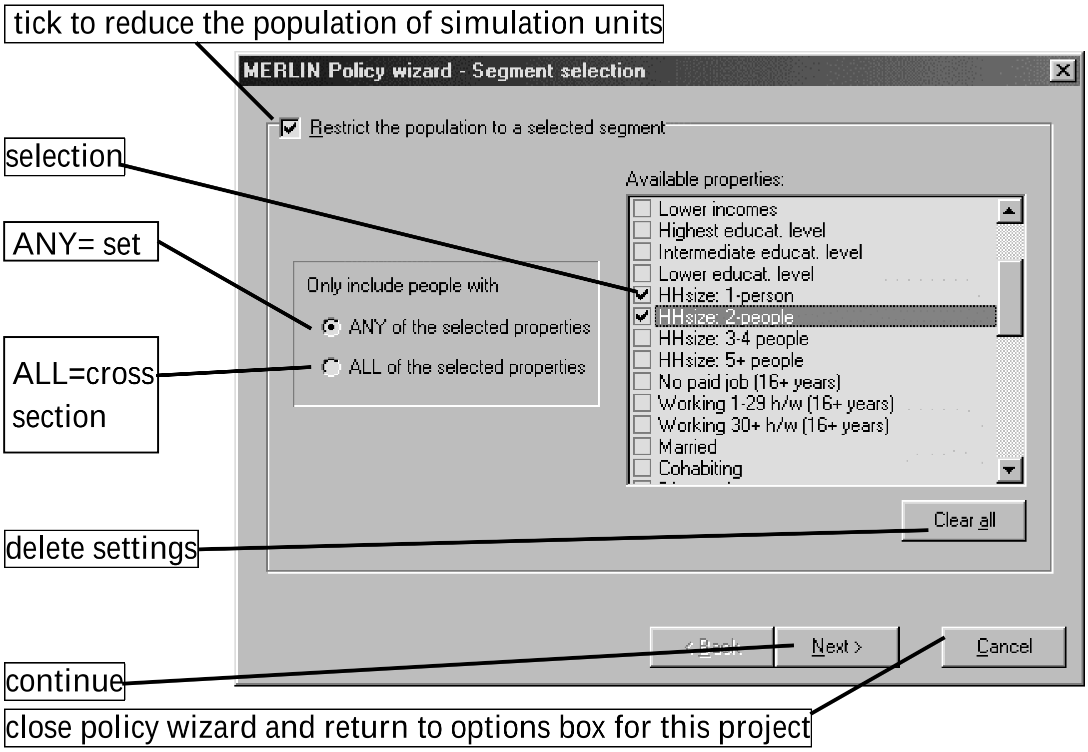


Figure A4.4 Dialogue box for selecting segments of the population

By default, none of the reallocation scenarios is activated. If one or more of these reallocation scenario's is (are) selected, the population of simulation units is (further) reduced to the relevant socio-economic group(s). The dialogue box (see Figure A4.6) allows you to specify the percentage of simulation units (belonging to the relevant socio-economic group) that has to be subjected to each particular reallocation scenario (except for scenario I, because this scenario applies to all students). If a particular reallocation scenario has been activated, its tab-leave is marked to signal this. The activation of reallocation scenarios can be undone by clicking "Disable all", which will deactivate the reallocation scenarios at once. Again, clicking "Cancel", will close down the policy wizard to return to the options box without storing your adjustments. Moving back and forth within the policy wizard by using the "< Back" and "Next >"-buttons will not delete your settings.

The next dialogue box of the policy wizard (Figure A4.7) allows you to change the aggregate preferences for transport modes (see also chapter 10 and section 12.4.2).

WARNING: Chapter 12 concluded that changing the aggregate preferences for transportation modes does not have the expected effects and should therefore not be used for impact analysis; see chapter 12 for details.

Group 1: People working at least 30 hours per week (in 1998)
Scenario A: Due to new collective labour agreements you had had 12 extra free days; your salary had been the same;
Scenario B: Due to new collective labour agreements you had had 12 free days less; your salary, however, had increased by 5%;
Scenario E: Due to new collective labour agreements you had worked 4 days a week; Friday had been your fixed free day; your salary had been the same.

Group 2: People working less than 30 h.p.w. with a partner working at least 30 h.p.w.
Scenario F: Due to new collective labour agreements your partner had had 12 extra free days; his/her salary had been the same;
Scenario G: Due to new collective labour agreements your partner had had 12 free days less; his/her salary, however, had been increased by 5 %.

Group 3: Students and pupils
Scenario I: In 1998 you had been allowed to use your Student Public Transport Ticket as it was introduced in 1991, viz. free use of all public transport modes within the Netherlands on all days.

Group 4: Pensioners, early retirement (Dutch: VUT) and people of independent means
Scenario J: Due to changes in the Dutch tax and social legislation, your net household income had been 15% higher in 1998;

Figure A4.5 The re-allocation scenarios available in the MERLIN-system

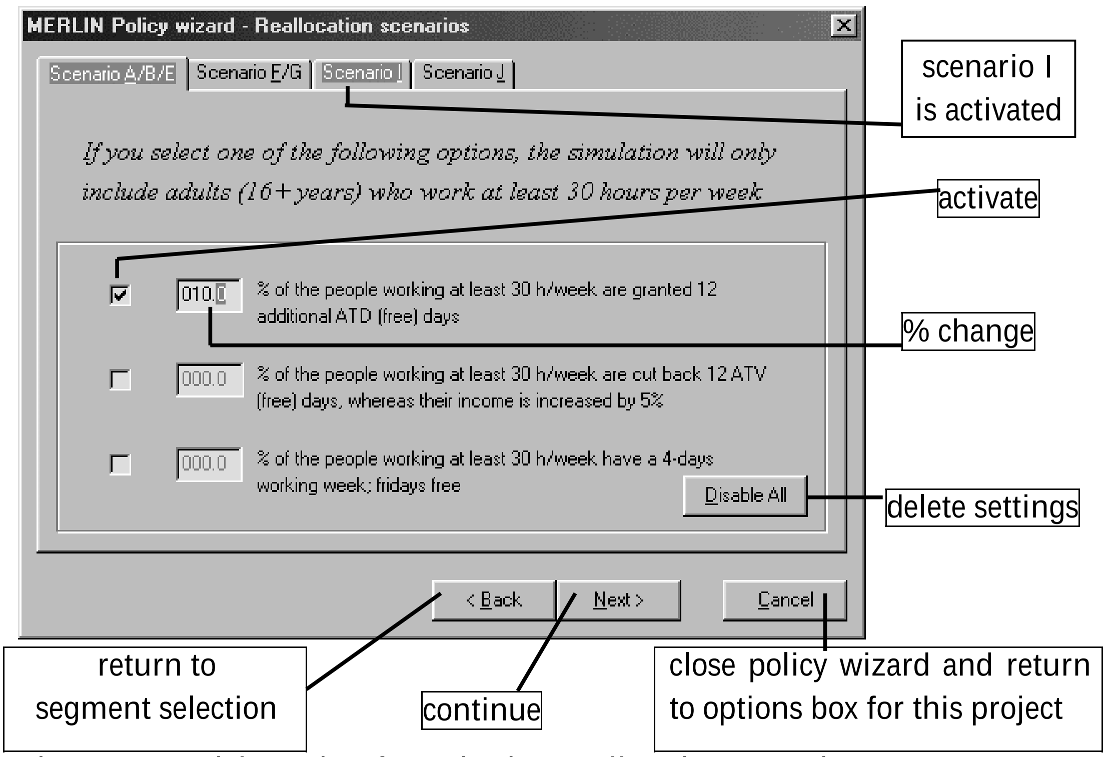


Figure A4.6 Dialogue box for activating re-allocation scenarios

By default, these preferences are 74.9% for car-based holidays, 12.2% for holidays by air and 12.9% for alternative means of transport. If you activate this type of changes for the project, these default values are activated and you can change them to any number within the range particular to that mode of transport. You will receive a warning if you try to continue to the next dialogue box while (1) the total percentage of aggregate preferences does not add up to 100%; and/or (2) at least one of values is outside the relevant range.

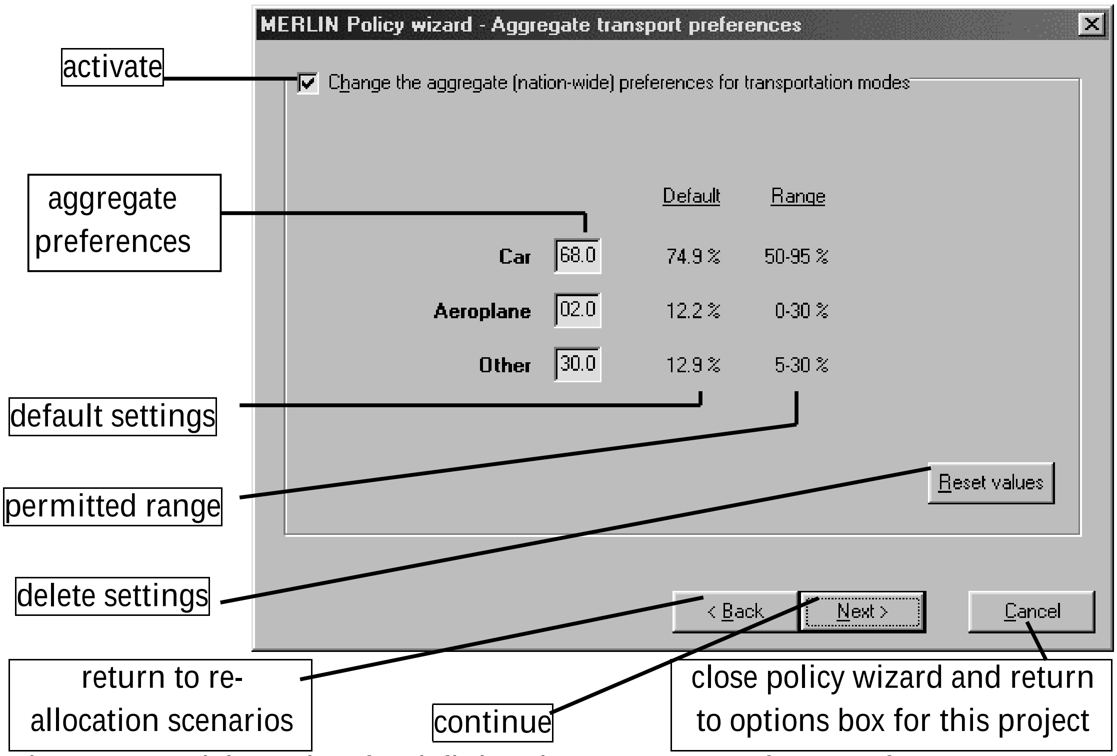


Figure A4.7 Dialogue box for defining the aggregate preferences for transport modes

The final dialogue box (see figure A4.8) of the policy wizard allows you to statically age and/or uprate the population. Characteristics of the population that can be uprated, including the possession of cars, skis, tourist accommodation and ‘being tied to holiday periods’, are indicated on the tab-leaves by the Roman numerals I, II, III and IV. Alternatively, changes to population regarding the age structure, education or income levels, the working situation, the presence of children and the household size, are implemented by statically ageing the population and are indicated in the policy wizard by the Arabic numerals (1-6).

If a particular characteristic of the population is activated, the *default settings* of this population attribute are transferred to the windows that can be adjusted. These default settings represent the proportion of simulation units in the selected population that meet particular characteristics. The *size of the selected population* (i.e. the combined result of the segment selection (Figure A4.4) and/or the selected reallocation scenario group (Figure A4.6)) is presented in the lower left corner of the dialogue box (WARNING: always check this size to make sure you did not specify an empty set!). Again, in this dialogue box it is important to make sure that the sum of percentages on each tab-leave adds up to 100% and that the individual values are all within the relevant ranges. Once you have completed the changes you wanted to implement, the “Finish”-button closes the policy wizard to return to the options box while storing the settings. Based on these settings, the specified population of simulation units can be simulated to assess the impact of the defined scenario for the future.

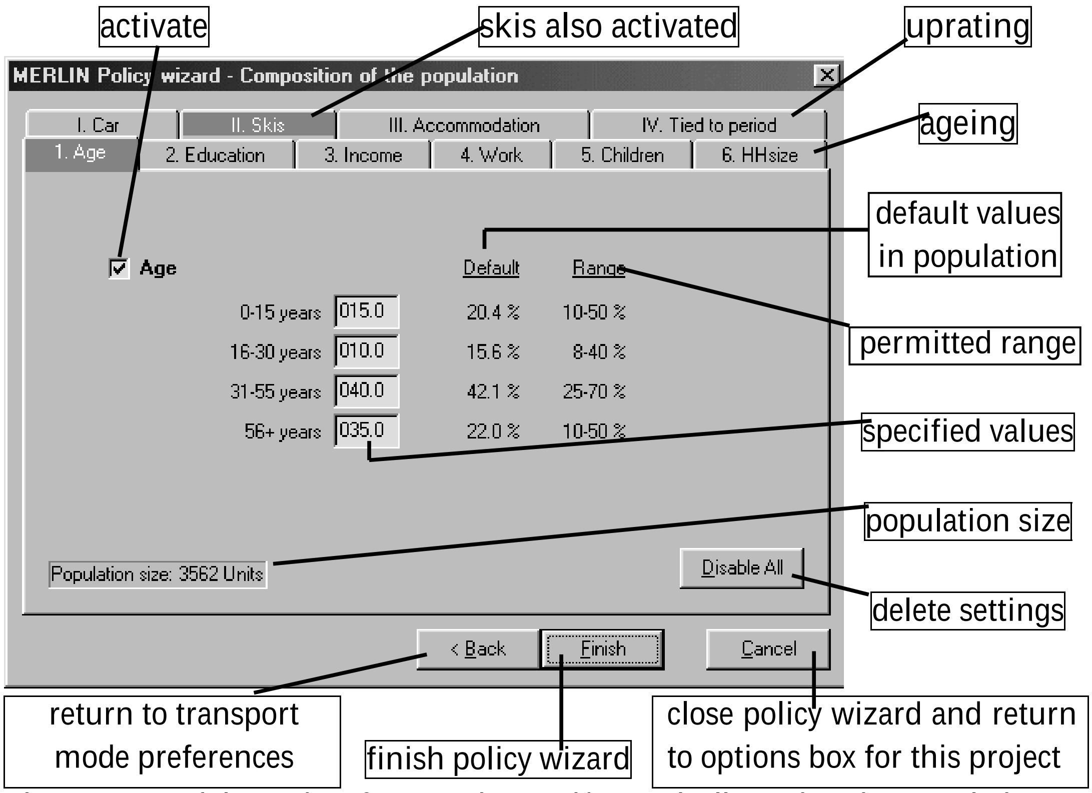
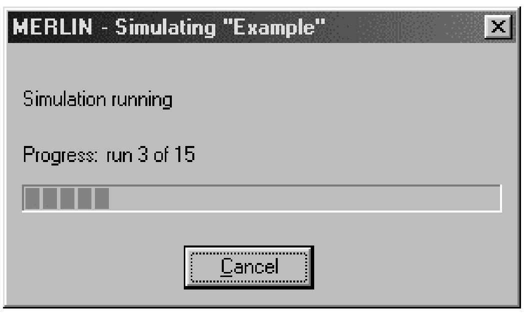


Figure A4.8 Dialogue box for uprating and/or statically ageing the population

A4.3 Viewing and interpreting results



Once the project and policy settings have been defined, the specified scenario for the future can be simulated by clicking the “Simulate”-button in the dialogue box for project options (Figure A4.3). The *MERLIN*-system will now run the simulation process *N* times, store the simulation results in the two databases (if this option is not disabled), and produce a standard output file (*.res) comprising several indicators of the simulation results at the aggregate level. This process can be interrupted at all times by clicking the “Cancel” button. If the simulation process is user-interrupted, the databases and the standard result file will be lost. The project and simulation settings, however, are still available for this project (stored in the *.pro file)[46].

A4.3.1 The aggregate results file

The *standard output file* of each successfully simulated project (*.res) is stored in the results directory. It is a text-file that can be viewed best using Microsoft Excel. To import the file:

1. Open Excel and click the “Open”-button
2. Browse to find the *.res file in the ‘\Merlin\Results’ directory using Files of type: All files (*.*)
3. Open the *.res file; Excel will now open the “Text Import Wizard”
4. In “Step 1 out of 3”, choose “Original data type” delimited; “Start import at row 1”; “File origin” Windows (ANSI); and click “Next >”
5. In “Step 2 out of 3”, choose “Delimiters” Tab; do NOT treat consecutive delimiters as one; “Text qualifier”; and click “Next >”
6. In “Step 3 out of 3”, select *all* imported text columns (scroll down and to the right and select all columns using the “shift”-button, choose “Column data format” → Text, and click “Finish” to finish the Wizard; Excel will now import the *.res file
7. Save the file as a Microsoft Excel Workbook (*.xls) file for further use.

The standard output file provides the following information:

(1) the population of simulation units, i.e., the input data, including information on income and educational levels, gender, school holiday region, the province of residence, civil status, age-categories, household size, presence of children and the number of people with paid jobs, car(s) in the household, permanent tourist accommodation, non-permanent tourist accommodation and skis;

(2) participation choices of the simulation units, including for both adults and

[46] WARNING: If you have interrupted the simulation process of an existing project (that has been simulated before) this means that the policy settings and the databases and aggregate results may not be related anymore while they do carry the same name!

children and itemised according to the participation choice groups, the average numbers of day-trips, short breaks, medium long holidays, extended holidays, long holidays, extra long holidays and days not allocated to tourism (per year per unit); and

(3) holiday choices of the simulation units, including the average duration of holidays (in days), the travel party, the timing (including season and school holiday periods), the destination, the accommodation, the means of transport, the day of departure, the average expenditures per (domestic or foreign) holiday and the average amount of total expenditures on holidays per year per income category.

Table A4.1 Population of Simulation Units according to age for "Example"

	1998 CVO Data	Reference		Example	
Age, Av. N° of					
- 0-15 year olds	726.00	726.00	(0.00)	532.67	(8.48)
- 16-30 year olds	555.00	555.00	(0.00)	352.67	(10.47)
- 31-55 year olds	1,498.00	1,498.00	(0.00)	1,423.07	(9.01)
- 56+ year olds	783.00	783.00	(0.00)	1,249.60	(15.44)
Average number of units (per run)	3,562.00	3,562.00	(0.00)	3,558.00	(22.33)

Tables A4.1 and A4.2 show some excerpts of the standard output file of the project "Example" that changed the age-composition of the population as shown in Figure A4.8 (all member of the 1998 CVO panel included; no policy settings other than the age-composition; N = 15). With regard to the *composition of the population* (Table A4.1), the standard output file shows:

- a label identifying a particular attribute (1st column);
- the number of respondents with the particular attribute in the 1998 CVO panel (2nd column);
- the average (across N runs) number of units with the particular attribute in the reference project (3rd column);
- the standard deviation of the average over N runs of the reference project (in brackets in the 4th column);
- the average number (across N runs) of units with the particular attribute in the current project (5th column);
- the standard deviation of the average over N runs of the current project (in brackets in the 6th column).

Table A4.2 Participation and holiday choices for “Example”

For all su's, the average N° of (per unit):	1998 CVO Data	Reference :		Example						
- DT (adults only)	14.9444	14.7403	(0.215)	14.1282	(0.136)	-0.612	-4.15%	-14.8564	26	**
- Holidays	1.9477	1.9971	(0.018)	1.9906	(0.021)	-0.007	-0.33%	-1.1685	17	
- SB	0.6804	0.7071	(0.009)	0.6744	(0.009)	-0.033	-4.63%	-13.3875	18	**
- MLH	0.5501	0.5684	(0.007)	0.5909	(0.010)	0.022	3.96%	8.1513	16	**
- EH	0.3932	0.4037	(0.008)	0.3991	(0.005)	-0.005	-1.14%	-2.9811	25	**
- LH	0.2872	0.2689	(0.007)	0.2637	(0.007)	-0.005	-1.91%	-2.7725	18	*
- ELH	0.0368	0.0490	(0.004)	0.0624	(0.003)	0.013	27.40%	15.6115	20	**
- days not allocated to tourism (adults)	331.4013	331.5207	(0.387)	331.9191	(0.311)	0.398	0.12%	4.4716	21	**
- days not allocated to tourism (children)	348.3072	348.2704	(0.513)	348.2531	(0.595)	-0.017	0.00%	-0.1071	17	
- Day of departure, Av. N° of holidays departing on										
- Fri/Sat	4,054.00	3,919.50	(61.20)	3,836.07	(72.08)	-83.433	-2.13%	-4.2587	17	**
- another day	3,067.00	3,194.19	(48.60)	3,246.53	(46.01)	52.343	1.64%	4.0780	19	**

In addition to the above columns, the standard output file also includes a comparison between the current and the reference project (see Table A4.2). Hence, in addition to the first six columns, these parts of the file also include:

- the absolute difference between the averages of the current and the reference project (7th column);
- the percentage difference between the averages of the current and the reference project (where the current project is the base; 8th column);
- the *t*-value of this difference (9th column);
- the degrees of freedom corrected for possible differences of standard deviations between the current and the reference project (10th column); and
- significance of the differences (2-sided test), where * indicates a significant difference at the 5% level of confidence, and ** indicates a significant difference at the 1% level of confidence.

From Table A4.1 it can be learned that the age-composition indeed changed according to the user-defined policy settings: on a total (average) population of 3558 units, the population included 532.67 (15.0%) units under 16 years; 352.67 (9.9%) units between 16 and 30 years; 1,423.07 (40.0%) units between 31 and 55 years; and 1,249.60 (35.1%) units aged 56 or older. Table A4.2 shows a significant increase in the average number of extra long holidays due to the increased share of elderly in the population in the "Example" project. Also, the number of holidays starting on Friday or Saturday decreases significantly (2.13%) due to this change.

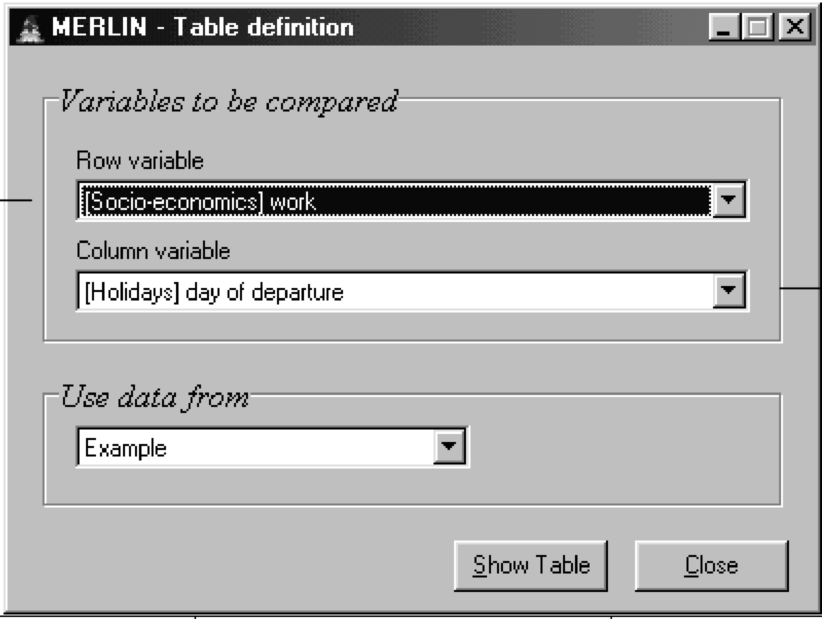
A4.3.2 The "Tables"-function

The second way to view and interpret the simulation results is to generate cross-tables that show the average frequencies of combinations of indicators. This option, activated by the "Tables"-button in the dialogue box for project options (Figure A4.3), can only be used when databases have been produced during the simulation process.

There are three types of variables to generate cross-tables, including socio-economics (i.e., the population), participation choices and holiday choices. Figure A4.9 shows the available variables for these types. Continuous variables are only available as row-variables. In this case, the table represents the average mean value and standard deviation of the continuous variable per category of the column value (including the average number of observations of these categories).

The combination of variable types in the cross-table determines the type of questions you can answer with regard to the simulation results. Table A4.3 provides an overview of interpretations for each combination of variable types.

Cross-tabulating two holiday-variables, for instance, will tell you something about the relationship between these two holiday choice facets. E.g., if people go on a holiday abroad are they more or less likely to leave on a Friday or Saturday compared to domestic trips? Or, with regard to the relationship between participation and holiday choices, do people who make more holidays tend to go abroad more often?



Socio-economics	Participation choices	Holidays
Income (8 categories)	Participation choices for adults	Duration (in days; 5 cats.)
Education (9 cats.)	Participation choices for children (0-15 years)	Travel party
Gender	# day-trips (adults only)*	In school holiday period or not
Region	# short breaks *	Season
Province	# medium long holidays *	Domestic or abroad
Age (4 categories)	# extended holidays *	Destination abroad (11 country zones)
Household size (5 cats.)	# long holidays *	Domestic destination (9 regions)
Children (3 cats.)	# extra long holidays *	Accommodation (4 cats.)
Work		Travel mode (5 cats.)
Car		Day of departure
Permanent tourist acc.		Expenditures per holiday *
Non-permanent tourist acc.		
Ski		
Annual expenditures *		

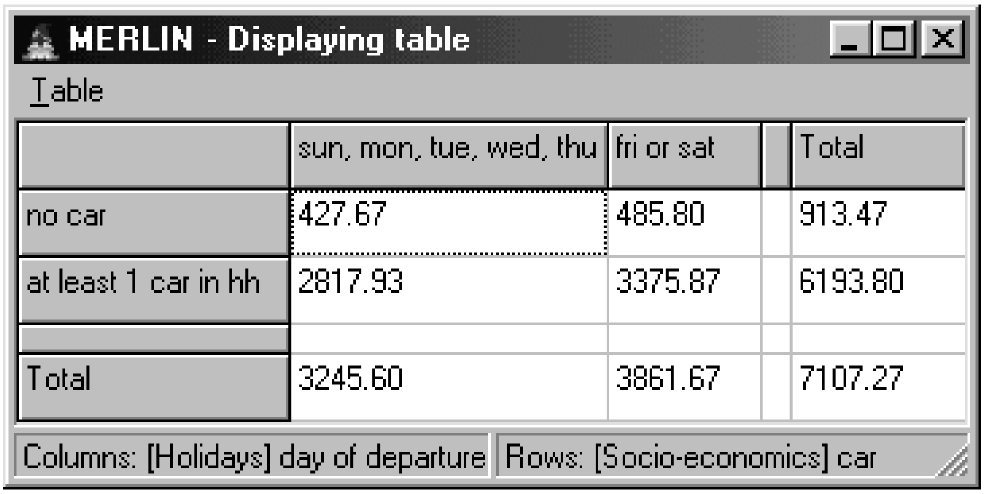
Figure A4.9 Variables available for cross-tables (variables marked by * are continuous and available as row variables only)

The final option in the "Table definition"-dialogue box is the selection of the data source. By default, the current project will be selected (see Figure A4.9). However, if you have installed the databases for the default reference project, you can also generate cross-tables from this source. In this case, select "1998 Predicted Data" from the "Use data from"-pull down menu (see Figure A4.9).

Table A4.3 Interpretation of cross-tables dependent on the type of variable

	Socio-economics (population)	Participation choices	Holidays
Socio-economics	Inspection of the composition of the population of simulation units	Relations between population segments with regard to participation in tourism	Relations between population segments with regard to facets of holiday choices
Participation choices		Relations between trip types (based on duration) within annual trip programs	Relations between annual trip choices and holiday-related choices
Holidays			Relations between facets of holidays

Once the row and column variables and the data source have been selected, click “Show Table” to generate the cross-table. The generated cross-table (see Figure A4.10) can be saved as a text file by selecting “Save” from the “Table” menu on the upper-left part of “Displaying table” box. You will be asked to enter a name, and by clicking “Save”, your table will be saved into the results directory with the default extension *.tab. For presentation purposes, this text-files can opened in a text-editor (in Microsoft Word, for instance, the text can easily be converted to a table by selecting the text in your text editor and selecting the “Convert Text to Table...”-option from the Table menu; tabs are used to separate the columns). Alternatively, the text file can be imported into spreadsheet applications like Microsoft Excel for further analysis.


MERLIN - Displaying table

Table

	sun, mon, tue, wed, thu	fri or sat		Total
no car	427.67	485.80		913.47
at least 1 car in hh	2817.93	3375.87		6193.80
Total	3245.60	3861.67		7107.27

Columns: [Holidays] day of departure | Rows: [Socio-economics] car

Figure A4.10 Generated cross-table of two variables in project “Example”

APPENDIX 5: Validation Results

This appendix presents the standard output file (slightly adjusted with regard to the presentation format) of the stability test. In this test, the results of the default reference project "*MERLIN*-null" are compared to the results of an identical simulation experiment (*N* = 200 runs, all members of the 1998 CVO-panel, without ageing and/or uprating). In addition, the standard output file comprises the observed tourist choices of the 1998 CVO-panel. By comparing the predicted tourist trip patterns of the 1998 CVO-panel ("*M*-null") to the observed patterns ("98 CVO"), the presence and extent of system irregularities at the aggregate level can be assessed.

Notes to the tables:

98 CVO — Observed tourist trip patterns of the members of the 1998 CVO-panel

M-null — Standard reference project "*MERLIN*-null", i.e. the predicted tourist trip patterns of the members of the 1998 CVO-panel (*N*=200), including the standard deviation (in brackets)

Test — Predicted tourist trip patterns of the members of the 1998 CVO-panel (*N*=200) (standard deviation in brackets) that are compared to "*MERLIN*-null" in order to assess the stability of the simulation results

diff. — Absolute difference between *M*-null and Test

%diff — Difference between *M*-null and Test as a % of *M*-null; apparent differences between "diff." and "%diff" are related to rounding off and the presentation format; calculations always use the exact figures

t-value — *t*-value of the difference between *M*-null and Test; in formula:

$$t = \frac{m_1 - m_2}{\sqrt{q_1 - q_2}} \text{ with } q_1 = \frac{s_1^{\ 2}}{n_1} \qquad and \qquad q_2 = \frac{s_2^{\ 2}}{n_2}$$

where m_1 , s_1^2 and n_1 are respectively the mean value, the variance and the number of runs for Test; and m_2, s_2^2 and n_2 are respectively the mean value, the variance and the number of runs for *M*-null df degrees of freedom of the *t*-value; in formula (all elements as before):

$$df = \frac{(q_1 + q_2)^2}{\frac{q_1^2}{n_1 - 1} + \frac{q_2^2}{n_2 - 1}}$$

S Significance of the *t*-test: '*' indicates a significant difference between *M*-null and Test at the 5% level (2-sided); and '**' indicates a significant difference between *M*-null and Test at the 1% level (2-sided)

The following tables compare the results of the default reference project "*MERCIN*-null" to the results of an identical simulation experiment (*N* = 200 runs, all members of the 1998 CVO-panel, without ageing and/or uprating). In addition, the tables comprise the observed tourist choices of the 1998 CVO-panel.

Table A5.1 Standard output file for the stability test – Participation choices

	98 CVO	*M*-null	Test	diff.	%diff	*t*-value	df	S
Av. N° of adults (16+ years):	1,530.00	2,836.00 (0.00)	2,836.00 (0.00)					
Av. N° of children (0-15 years):	726.00	726.00 (0.00)	726.00 (0.00)					

Table A5.1 Participation choices - Continued

	98 CVO	$\mathcal{M}$-null	Test	diff.	%diff	*t*-value	df	S
For all simulation units, the average N° of (per unit):								
- Day-trips (DT) (adults only)	14.9444	14.6521 (0.163)	14.6417 (0.167)	-0.010	-0.07%	-0.6307	398	
- Holidays	1.9477	2.0009 (0.021)	1.9999 (0.019)	-0.001	-0.05%	-0.5021	393	
- Short Breaks (SB)	0.6804	0.7082 (0.009)	0.7075 (0.009)	-0.001	-0.10%	-0.8082	398	
- Medium Long Holidays (MLH)	0.5501	0.5694 (0.008)	0.5697 (0.008)	0.000	0.04%	0.2537	396	
- Extended Holidays (EH)	0.3932	0.4038 (0.008)	0.4038 (0.009)	0.000	0.01%	0.0562	397	
- Long Holidays (LH)	0.2872	0.2705 (0.007)	0.2701 (0.007)	0.000	-0.16%	-0.6339	398	
- Extra Long Holidays (ELH)	0.0368	0.0489 (0.003)	0.0489 (0.003)	0.000	-0.17%	-0.2719	396	
- days not allocated to tourism (adults)	331.4013	331.5401 (0.368)	331.5630 (0.301)	0.023	0.01%	0.6818	383	
- days not allocated to tourism (children)	348.3072	348.3092 (0.498)	348.3206 (0.469)	0.011	0.00%	0.2370	397	
Av. N° of adults without tourist trips	195.00	361.05 (16.19)	362.04 (14.42)	0.990	0.27%	0.6458	393	
Av. N° of children (0-15 years) without tourist trips	114.00	113.33 (10.19)	115.02 (9.49)	1.685	1.49%	1.7112	396	
Av. N° of adults making day-trips only	133.00	251.68 (15.32)	251.88 (13.21)	0.205	0.08%	0.1433	390	
- Av. N° of day-trips per unit	20.2256	19.3028 (0.527)	19.2968 (0.536)	-0.006	-0.03%	-0.1135	398	
Av. N° of adults going on holidays only	259.00	477.70 (18.36)	478.39 (19.70)	0.690	0.14%	0.3624	396	
- Av. N° of holidays per unit	2.3359	2.4091 (0.082)	2.4028 (0.080)	-0.006	-0.26%	-0.7722	398	
- Av. N° of SB per unit	0.7452	0.8127 (0.060)	0.8063 (0.053)	-0.006	-0.79%	-1.1284	392	
- Av. N° of MLH per unit	0.6718	0.6679 (0.025)	0.6687 (0.024)	0.001	0.11%	0.2946	397	
- Av. N° of EH per unit	0.4749	0.4884 (0.024)	0.4866 (0.021)	-0.002	-0.36%	-0.7788	390	
- Av. N° of LH per unit	0.3861	0.3541 (0.022)	0.3534 (0.022)	-0.001	-0.18%	-0.2895	398	
- Av. N° of ELH per unit	0.0579	0.0860 (0.012)	0.0879 (0.013)	0.002	2.16%	1.4537	396	
- Av. N° of days not allocated to tourism per unit	342.1313	341.6551 (0.754)	341.6489 (0.839)	-0.006	0.00%	-0.0769	394	

Table A5.1 Participation choices - Continued

	98 CVO	$\mathcal{M}$-null	Test	diff.	%diff	*t*-value	df	S
Av. N° of children (0-15 years) going on holidays	612.00	612.67 (10.19)	610.99 (9.49)	-1.685	-0.28%	-1.7112	396	
- Av. N° of holidays per unit	2.1454	2.1387 (0.039)	2.1459 (0.038)	0.007	0.33%	1.8479	398	
- Av. N° of SB per unit	0.7500	0.7457 (0.020)	0.7497 (0.020)	0.004	0.55%	2.0528	398	*
- Av. N° of MLH per unit	0.5980	0.5985 (0.018)	0.5999 (0.020)	0.001	0.25%	0.7846	396	
- Av. N° of EH per unit	0.4363	0.4355 (0.021)	0.4365 (0.021)	0.001	0.22%	0.4640	398	
- Av. N° of LH per unit	0.3333	0.3308 (0.017)	0.3321 (0.018)	0.001	0.41%	0.7724	396	
- Av. N° of ELH per unit	0.0278	0.0283 (0.007)	0.0275 (0.006)	-0.001	-2.74%	-1.2023	391	
- Av. N° of days not allocated to tourism per unit	345.1977	345.2209 (0.524)	345.1799 (0.504)	-0.041	-0.01%	-0.7984	397	
Av. N° of adults going on holidays & day-trips	943.00	1,745.57 (21.52)	1,743.68 (22.08)	-1.885	-0.11%	-0.8646	398	
- Av. N° of day-trips per unit	21.3945	21.0224 (0.090)	21.0268 (0.084)	0.004	0.02%	0.5105	396	
- Av. N° of holidays per unit	2.6257	2.6732 (0.029)	2.6745 (0.027)	0.001	0.05%	0.4629	395	
- Av. N° of SB per unit	0.9364	0.9611 (0.016)	0.9614 (0.014)	0.000	0.03%	0.2075	393	
- Av. N° of MLH per unit	0.7434	0.7691 (0.012)	0.7700 (0.012)	0.001	0.11%	0.7600	398	
- Av. N° of EH per unit	0.5270	0.5375 (0.011)	0.5385 (0.012)	0.001	0.19%	0.8732	395	
- Av. N° of LH per unit	0.3648	0.3391 (0.011)	0.3384 (0.011)	-0.001	-0.19%	-0.5907	398	
- Av. N° of ELH per unit	0.0541	0.0664 (0.005)	0.0661 (0.005)	0.000	-0.48%	-0.6127	398	
- Av. N° of days not allocated to tourism per unit	319.6204	319.8098 (0.388)	319.8092 (0.360)	-0.001	0.00%	-0.0167	396	

Table A5.2 Standard output file for the stability test – Holiday choices

	98 CVO	$\mathcal{M}$-null	Test	diff.	%diff	*t*-value	df	S
Av. duration of holidays (in days)	9.18	9.18 (0.06)	9.18 (0.06)	-0.001	-0.02%	-0.2238	396	
Travel party, Av. N° of holidays with								
- alone	248.00	288.26 (17.91)	288.68 (15.25)	0.410	0.14%	0.2465	388	
- adults only	2,707.00	2,756.85 (47.18)	2,754.76 (47.90)	-2.095	-0.08%	-0.4406	398	
- schoolchildren (6-14 years)	1,370.00	1,425.93 (36.26)	1,424.76 (34.43)	-1.175	-0.08%	-0.3323	397	
- other children (0-5 or 15+ years)	933.00	981.63 (28.88)	982.30 (30.07)	0.670	0.07%	0.2272	397	
- party of 9+ people	717.00	750.48 (27.26)	749.83 (25.64)	-0.645	-0.09%	-0.2438	397	
- unknown	1,146.00	924.18 (29.66)	923.43 (29.78)	-0.750	-0.08%	-0.2523	398	
Season & In-/outside school holiday period, Av. N° of holidays in								
- winter, during school holiday period	399.00	435.61 (20.73)	435.02 (20.36)	-0.585	-0.13%	-0.2847	398	
- winter, outside school holiday period	447.00	464.72 (24.25)	466.65 (20.50)	1.930	0.42%	0.8595	387	
- spring, during school holiday period	944.00	1,028.88 (31.04)	1,024.23 (29.80)	-4.645	-0.45%	-1.5266	397	
- spring, outside school holiday period	1,607.00	1,563.60 (38.58)	1,561.40 (38.06)	-2.200	-0.14%	-0.5741	398	
- summer-begin	878.00	832.73 (23.57)	827.16 (25.53)	-5.565	-0.67%	-2.2648	395	*
- summer-mid	1,082.00	1,059.38 (31.54)	1,060.25 (29.05)	0.870	0.08%	0.2869	395	
- summer-end	652.00	637.14 (25.35)	642.58 (23.47)	5.430	0.85%	2.2231	396	*
- autumn, during school holiday period	322.00	334.27 (19.68)	335.13 (18.64)	0.860	0.26%	0.4487	397	
- autumn, outside school holiday period	790.00	771.00 (27.25)	771.32 (24.96)	0.320	0.04%	0.1224	395	
Destination, Av. N° of holidays to/in:								
* The Netherlands	3,855.00	3,735.30 (60.82)	3,731.28 (58.60)	-4.030	-0.11%	-0.6748	397	
♠ metaregion "Water"	1,177.00	1,188.24 (39.98)	1,189.67 (43.48)	1.430	0.12%	0.3424	395	
- North	469.00	454.77 (22.97)	453.79 (25.18)	-0.985	-0.22%	-0.4087	395	
- Mid	177.00	187.31 (18.93)	187.24 (21.60)	-0.075	-0.04%	-0.0369	391	
- North Sea	531.00	546.16 (28.99)	548.65 (29.14)	2.490	0.46%	0.8567	398	

Table A5.2 Holiday choices - Continued

	98 CVO	$\mathcal{M}$-null	Test	diff.	%diff	*t*-value	df	S
♠ metaregion "Land north"	1,491.00	1,458.83 (47.20)	1,457.50 (46.46)	-1.325	-0.09%	-0.2829	398	
- Mid	580.00	549.89 (29.81)	549.17 (28.76)	-0.725	-0.13%	-0.2475	397	
- East	485.00	527.16 (32.45)	525.38 (33.59)	-1.790	-0.34%	-0.5420	398	
- North	426.00	381.76 (25.84)	382.95 (25.90)	1.190	0.31%	0.4600	398	
♠ metaregion "Land south"	976.00	877.61 (37.36)	872.15 (37.00)	-5.455	-0.62%	-1.4671	398	
- West	465.00	287.60 (22.97)	285.60 (25.18)	-2.005	-0.70%	-0.8547	398	
- South	511.00	590.01 (18.93)	586.56 (21.60)	-3.450	-0.58%	-1.2485	395	
♠ metaregion "Other"	211.00	210.63 (20.55)	211.94 (18.46)	1.320	0.63%	0.6758	393	
* Foreign countries	3,266.00	3,392.03 (59.21)	3,392.47 (56.66)	0.445	0.01%	0.0768	397	
♣ neighbouring countries	1,656.00	1,827.89 (42.82)	1,828.56 (41.40)	0.665	0.04%	0.1579	398	
- France	656.00	706.58 (25.59)	703.70 (24.00)	-2.885	-0.41%	-1.1629	396	
- Belgium & Luxembourg	423.00	474.44 (21.32)	477.73 (20.74)	3.290	0.69%	1.5643	398	
- Germany	392.00	445.20 (21.80)	446.53 (23.18)	1.325	0.30%	0.5889	397	
- United Kingdom	185.00	201.68 (15.97)	200.61 (13.07)	-1.065	-0.53%	-0.7300	383	
♣ more distant countries	1,610.00	1,565.10 (41.69)	1,565.02 (39.24)	-0.085	-0.01%	-0.0210	397	
- Spain & Portugal	417.00	432.98 (20.23)	432.25 (20.21)	-0.725	-0.17%	-0.3586	398	
- Austria & Switzerland	322.00	273.07 (15.79)	273.56 (15.40)	0.490	0.18%	0.3142	398	
- Italy & Greece	263.00	282.76 (16.02)	283.03 (17.75)	0.265	0.09%	0.1568	394	
- Eastern Europe	141.00	150.38 (11.74)	149.34 (12.48)	-1.035	-0.69%	-0.8540	397	
- South East Mediterranean	116.00	117.75 (10.34)	119.51 (11.39)	1.755	1.49%	1.6127	394	
- Scandinavia	93.00	99.11 (10.15)	98.41 (9.58)	-0.695	-0.70%	-0.7040	397	
- Other	258.00	208.08 (13.64)	207.81 (13.35)	-0.275	-0.13%	-0.2037	398	

Table A5.2 Holiday choices - Continued

	98 CVO	$\mathcal{M}$-null	Test	diff.	%diff	*t*-value	df	S
Accommodation, Av. N° of holidays in								
- hotel	2,347.00	2,440.68	2,438.70	-1.975	-0.08%	-0.3734	398	
		(53.30)	(52.49)					
- non-permanent-own	998.00	921.24	921.17	-0.070	-0.01%	-0.0232	397	
		(30.87)	(29.55)					
- permanent-own	1,170.00	975.52	975.88	0.365	0.04%	0.0989	398	
		(37.15)	(36.69)					
- other	2,606.00	2,789.89	2,787.99	-1.905	-0.07%	-0.3790	384	
		(54.92)	(45.14)					
Means of transport, Av. N° of holidays by								
- car	5,335.00	5,321.90	5,320.12	-1.785	-0.03%	-0.2991	392	
		(63.27)	(55.86)					
- aeroplane	868.00	812.95	814.09	1.140	0.14%	0.3986	397	
		(29.47)	(27.71)					
- alternative, of which	918.00	992.48	989.54	-2.940	-0.30%	-0.9172	398	
		(32.35)	(31.76)					
- bus	416.00	449.74	449.50	-0.240	-0.05%	-0.1136	398	
		(21.04)	(21.20)					
- train	260.00	274.26	274.12	-0.140	-0.05%	-0.0875	398	
		(16.16)	(15.83)					
- other	242.00	268.48	265.92	-2.560	-0.95%	-1.5157	397	
		(16.45)	(17.32)					
Day of departure, Av. N° of holidays departing on								
- Friday/Saturday	4,054.00	3,922.80	3,918.06	-4.745	-0.12%	-0.8112	393	
		(61.61)	(55.21)					
- other day of the week	3,067.00	3,204.53	3,205.69	1.160	0.04%	0.2232	398	
		(52.75)	(51.20)					
Average expenditures (NLG) per								
* holiday	713.03	744.50	744.45	-0.049	-0.01%	-0.0450	393	
		(11.51)	(10.32)					
♠ domestic holiday	268.42	289.04	288.41	-0.633	-0.22%	-1.2634	394	
		(4.75)	(5.26)					
- domestic SB	155.87	180.01	179.86	-0.152	-0.08%	-0.3756	395	
		(3.87)	(4.20)					
- domestic MLH	299.24	321.02	320.70	-0.327	-0.10%	-0.3882	398	
		(8.35)	(8.52)					
- domestic EH	468.70	491.47	490.33	-1.140	-0.23%	-0.5328	398	
		(21.47)	(21.31)					
- domestic LH	617.52	559.49	553.15	-6.338	-1.13%	-1.8348	396	
		(33.17)	(35.87)					
- domestic ELH	803.92	541.43	536.10	-5.330	-0.98%	-0.7384	395	
		(75.17)	(69.07)					

Table A5.2 Holiday choices - Continued

	98 CVO	$\mathcal{M}$-null	Test	diff.	%diff	*t*-value	df	S
♣ holiday abroad	1,237.82	1,246.08 (20.40)	1,246.07 (17.05)	-0.016	0.00%	-0.0083	386	
- SB abroad	386.51	507.23 (25.39)	508.18 (24.01)	0.948	0.19%	0.3838	397	
- MLH abroad	846.23	1,038.07 (35.43)	1,043.34 (33.11)	5.267	0.51%	1.5360	396	
- EH abroad	1,393.99	1,503.14 (40.25)	1,499.98 (36.09)	-3.152	-0.21%	-0.8246	393	
- LH abroad	1,761.05	1,685.43 (45.97)	1,681.58 (43.60)	-3.850	-0.23%	-0.8593	397	
- ELH abroad	3,242.39	2,043.39 (106.61)	2,061.05 (120.12)	17.666	0.86%	1.5556	392	

APPENDIX 6: Results of Demonstration Projects

This appendix presents the standard output file (slightly adjusted with regard to the presentation format) of the three demonstration projects. Each demonstration project, i.e. scenario for the future, is compared to the appropriate baseline simulation. In addition, the results of the test for significant differences for a second (similar) project are presented to exclude differences that are based on chance.

Notes to the tables:

Base	Baseline simulation; i.e. the predicted tourist trip patterns of (a particular segment of) the members of the 1998 CVO-panel, including the standard deviation (in brackets)
Demo *X*	Predicted tourist trip patterns of (a particular segment of) the members of the 1998 CVO-panel (*N*=25) (standard deviation in brackets) including the proposed scenario for the future (Demo A, Demo B and Demo C)
diff.	Absolute difference between the baseline and the scenario for the future
%diff	Difference between the baseline and the scenario for the future as a % of the baseline simulation; apparent differences between "diff." and "%diff" are related to rounding off and the presentation format; calculations always use the exact figures
t-value	*t*-value of the difference between the baseline simulation and the scenario for the future
df	degrees of freedom of the *t*-value
S1	Significance of the *t*-test: '*' indicates a significant difference between the baseline and the scenario for the future at the 5% significance level (2-sided); and '**' indicates a significant difference at the 1% significance level (2-sided)
S2	Results of the significance test for the baseline and the same scenario for the future (see also S1); differences between S1 and S2 indicate that the alleged significant change is likely to be based on chance; if the differences between the baseline simulation and the scenario for the future are significant for both scenario-projects, the cells are framed to mark this correspondence

A6.1 Demonstration A: 10% of the full-time working members of the population are granted 12 extra free days per year

Table A6.1 Standard output file for demonstration project A: 10% of the full-time working members of the population is granted 12 extra free days per year – Participation choices

	Base	Demo A	diff.	%diff	*t*-value	df	S1	S2
Av. N° of adults (16+ years):	1,192.00 (.00)	1,192.00 (.00)						
Av. N° of children (0-15 years):	.00 (.00)	.00 (.00)						
For all simulation units, the average N° of (per unit):								
- Day-trips (DT) (adults only)	15.7707 (.237)	15.7005 (.234)	-.070	-.45%	-1.0554	48		
- Holidays	2.1047 (.029)	2.1611 (.031)	.056	2.68%	6.5720	48	**	**
- Short Breaks (SB)	.7907 (.017)	.8241 (.017)	.033	4.21%	6.8957	48	**	**
- Medium Long Holidays (MLH)	.5675 (.013)	.5853 (.009)	.018	3.15%	5.7019	44	**	**
- Extended Holidays (EH)	.4471 (.013)	.4463 (.016)	-.001	-.17%	-.1869	46		
- Long Holidays (LH)	.2790 (.013)	0.2829 (0.011)	.004	1.38%	1.1308	47		
- Extra Long Holidays (ELH)	.0204 (.003)	.0225 (.004)	.002	10.18%	2.0442	38	*	*
- days not allocated to tourism (adults)	330.7688 (.460)	330.4760 (.387)	-.293	-.09%	-2.4348	47	*	
- days not allocated to tourism (children)	.0000 (.000)	.0000 (.000)						
Av. N° of adults without tourist trips	79.92 (7.16)	80.20 (7.26)	.28	.35%	.1373	48		
Av. N° of children without tourist trips	.0000 (.000)	.0000 (.000)						
Av. N° of adults making day-trips only	98.48 (9.74)	97.20 (10.74)	-1.28	-1.30%	-.4414	48		
- Av. N° of day-trips per unit	18.5971 (.710)	18.8129 (.544)	.216	1.16%	1.2069	45		

Table A6.1 Participation choices for demonstration project A- Continued

	Base	Demo A	diff.	%diff	*t*-value	df	S1	S2
Av. N° of adults going on holidays only	199.68 (11.70)	211.60 (13.31)	11.92	5.97%	3.3631	47	**	**
- Av. N° of holidays per unit	2.2048 (0.112)	2.2360 (.125)	.031	1.42%	.9308	47		
- Av. N° of SB per unit	.7013 (0.080)	.7240 (.077)	.023	3.24%	1.0208	48		
- Av. N° of MLH per unit	.6043 (.038)	.6449 (.034)	.041	6.71%	3.9717	47	**	**
- Av. N° of EH per unit	.5356 (.029)	.5224 (.037)	-.013	-2.47%	-1.3952	45		*
- Av. N° of LH per unit	.2772 (.034)	.2511 (.032)	-.026	-9.40%	-2.7751	48	**	*
- Av. N° of ELH per unit	.0864 (.015)	.0935 (.021)	.007	8.31%	1.3727	43		
- Av. N° of days not allocated to tourism per unit	343.5502 (1.275)	343.7146 (1.489)	.164	.05%	.4194	47		
Av. N° of adults going on holidays & day-trips	813.92 (14.52)	803.00 (15.09)	-10.92	-1.34%	-2.6072	48	*	
- Av. N° of day-trips per unit	20.8462 (.134)	21.0253 (.117)	.179	.86%	5.0218	47	**	**
- Av. N° of holidays per unit	2.5413 (.031)	2.6190 (.042)	.078	3.06%	7.4773	44	**	**
- Av. N° of SB per unit	.9857 (.022)	1.0323 (.025)	.047	4.73%	7.0037	47	**	**
- Av. N° of MLH per unit	.6829 (.014)	.6991 (.015)	.016	2.38%	3.9757	48	**	**
- Av. N° of EH per unit	.5233 (.017)	.5250 (.020)	.002	.33%	.3286	47		
- Av. N° of LH per unit	.3406 (.016)	.3538 (.015)	.013	3.88%	3.0128	48	**	*
- Av. N° of ELH per unit	.0088 (.003)	.0088 (.003)	.000	-.28%	-.0282	47		
- Av. N° of days not allocated to tourism per unit	322.3802 (.458)	321.6373 (.412)	-.743	-.23%	-6.0251	47	**	**

Table A6.2 Standard output file for demonstration project A: 10% of the full-time working members of the population is granted 12 extra free days per year – Holiday choices

	Base	Demo A	diff.	%diff	*t*-value	df	S1	S2
Av. duration of holidays (in days)	8.77 (.10)	8.71 (.07)	-.06	-.69%	-2.5337	45	*	

Table A6.2 Holiday choices for demonstration project A- Continued

	Base	Demo A	diff.	%diff	*t*-value	df	S1	S2
Travel party, Av. N° of holidays with								
- alone	147.52 (11.42)	146.64 (13.10)	-.88	-.60%	-.2532	47		
- adults only	1,224.32 (23.42)	1,249.20 (27.59)	24.88	2.03%	3.4375	47	**	**
- schoolchildren (6-14 years)	333.76 (18.03)	343.04 (16.17)	9.28	2.78%	1.9158	47		
- other children (0-5 or 15+ years)	286.12 (17.91)	300.92 (18.62)	14.80	5.17%	2.8644	48	**	**
- party of 9+ people	248.48 (17.65)	251.56 (16.33)	3.08	1.24%	.6405	48		
- unknown	268.64 (11.08)	284.64 (15.50)	16.00	5.96%	4.1995	43	**	**
Season & In-/outside school holiday period, Av. N° of holidays in								
- winter, during school holiday period	148.96 (10.46)	152.44 (10.66)	3.48	2.34%	1.1649	48		
- winter, outside school holiday period	178.32 (13.70)	180.92 (13.07)	2.60	1.46%	.6866	48		
- spring, during school holiday period	336.40 (14.82)	351.16 (21.18)	14.76	4.39%	2.8549	43	**	**
- spring, outside school holiday period	553.36 (19.11)	568.88 (24.02)	15.52	2.80%	2.5283	46	*	*
- summer-begin	280.28 (13.46)	287.48 (16.25)	7.20	2.57%	1.7063	46		
- summer-mid	366.44 (22.27)	368.60 (16.11)	2.16	.59%	.3930	44		
- summer-end	228.84 (12.25)	235.20 (14.61)	6.36	2.78%	1.6674	47		
- autumn, during school holiday period	115.68 (9.28)	120.36 (14.36)	4.68	4.05%	1.3687	41		
- autumn, outside school holiday period	300.56 (20.00)	310.96 (18.80)	10.40	3.46%	1.8943	48		
Destination, Av. N° of holidays to/in:								
* The Netherlands	1,136.36 (40.56)	1,172.20 (29.00)	35.84	3.15%	3.5940	43	**	*
♠ metaregion "Water"	373.32 (19.71)	383.64 (26.03)	10.32	2.76%	1.5803	45		
- North	145.52 (13.34)	144.16 (13.50)	-1.36	-.93%	-.3583	48		
- Mid	55.52 (7.93)	58.84 (10.79)	3.32	5.98%	1.2394	44		
- North Sea	172.28 (14.22)	180.64 (19.67)	8.36	4.85%	1.7222	44		

Table A6.2 Holiday choices for demonstration project A- Continued

	Base	Demo A	diff.	%diff	*t*-value	Df	S1	S2
♠ metaregion "Land north"	424.64 (27.79)	450.16 (24.47)	25.52	6.01%	3.4459	47	**	*
- Mid	166.60 (16.38)	171.40 (18.31)	4.80	2.88%	.9770	47		
- East	151.16 (17.56)	158.04 (12.93)	6.88	4.55%	1.5774	44		*
- North	106.88 (16.45)	120.72 (16.11)	13.84	12.95%	3.0052	48	**	
♠ metaregion "Land south"	270.72 (20.70)	268.36 (26.71)	-2.36	-.87%	-.3492	45		
- West	90.92 (13.34)	87.28 (13.50)	-3.64	-4.00%	-.8351	46		
- South	179.80 (7.93)	181.08 (10.79)	1.28	.71%	.3029	47		
♠ metaregion "Other"	67.68 (10.42)	70.04 (11.69)	2.36	3.49%	.7535	47		
* Foreign countries	1,372.48 (35.19)	1,403.80 (32.95)	31.32	2.28%	3.2485	48	**	**
♣ neighbouring countries	720.28 (23.78)	741.24 (21.98)	20.96	2.91%	3.2367	48	**	**
- France	270.00 (17.73)	276.48 (14.36)	6.48	2.40%	1.4201	46		
- Belgium & Luxembourg	189.40 (11.92)	197.16 (10.87)	7.76	4.10%	2.4060	48	*	
- Germany	176.20 (12.96)	176.52 (9.84)	.32	.18%	.0983	45		
- United Kingdom	84.68 (8.55)	91.08 (9.89)	6.40	7.56%	2.4485	47	*	**
♣ more distant countries	652.20 (23.94)	662.56 (22.67)	10.36	1.59%	1.5712	48		
- Spain & Portugal	183.92 (15.08)	181.64 (13.71)	-2.28	-1.24%	-.5593	48		
- Austria & Switzerland	118.44 (9.90)	122.60 (9.54)	4.16	3.51%	1.5130	48		
- Italy & Greece	115.52 (12.09)	115.16 (12.89)	-.36	-.31%	-.1018	48		
- Eastern Europe	58.92 (8.50)	61.48 (7.82)	2.56	4.34%	1.1085	48		
- South East Mediterranean	42.80 (5.61)	45.28 (5.97)	2.48	5.79%	1.5144	48		
- Scandinavia	39.60 (6.34)	39.32 (6.38)	-.28	-.71%	-.1556	48		
- Other	93.00 (10.33)	97.08 (6.14)	4.08	4.39%	1.6976	39		*

Table A6.2 Holiday choices for demonstration project A- Continued

	Base	Demo A	diff.	%diff	*t*-value	df	S1	S2
Accommodation, Av. N° of holidays in								
- hotel	1,009.08 (32.10)	1,034.12 (31.93)	25.04	2.48%	2.7653	48	**	**
- non-permanent-own	320.08 (18.40)	326.52 (16.18)	6.44	2.01%	1.3139	47		
- permanent-own	268.64 (15.37)	282.20 (19.80)	13.56	5.05%	2.7046	45	**	
- other	911.04 (27.16)	933.16 (23.42)	22.12	2.43%	3.0840	47	**	**
Means of transport, Av. N° of holidays by								
- car	1,785.32 (42.22)	1,833.60 (33.32)	48.28	2.70%	4.4884	46	**	**
- aeroplane	360.12 (18.11)	368.56 (17.02)	8.44	2.34%	1.6979	48		
- alternative, of which	363.40 (16.79)	373.84 (16.26)	10.44	2.87%	2.2337	48	*	
- bus	169.00 (12.66)	174.80 (11.74)	5.80	3.43%	1.6799	48		
- train	94.56 (10.76)	97.84 (9.93)	3.28	3.47%	1.1200	48		
- other	99.84 (9.37)	101.20 (10.92)	1.36	1.36%	0.4724	47		
Day of departure, Av. N° of holidays departing on								
- Friday/Saturday	1,391.08 (28.09)	1,427.68 (32.66)	36.60	2.63%	4.2482	47	**	**
- other day of the week	1,117.76 (32.54)	1,148.32 (28.69)	30.56	2.73%	3.5219	47	**	**
Average expenditures (NLG) per								
* holiday	855.98 (21.49)	847.38 (19.04)	-8.60	-1.01%	-1.4988	47		
♠ domestic holiday	277.81 (8.37)	280.69 (8.13)	2.88	1.04%	1.2346	48		
- domestic SB	196.19 (7.02)	197.31 (5.59)	1.12	.57%	.6244	46		
- domestic MLH	328.53 (14.79)	330.96 (13.47)	2.42	.74%	.6056	48		
- domestic EH	479.75 (43.78)	489.44 (43.68)	9.69	2.02%	.7835	48		
- domestic LH	541.17 (57.13)	581.53 (84.79)	40.36	7.46%	1.9736	42		
- domestic ELH	357.02 (168.11)	477.19 (266.84)	120.18	33.66%	1.9053	40		*

Table A6.2 Holiday choices for demonstration project A- Continued

	Base	Demo A	diff.	%diff	*t*-value	df	S1	S2
♣ holiday abroad	1,334.82 (36.56)	1,320.53 (29.35)	-14.29	-1.07%	-1.5247	46		*
- SB abroad	518.62 (35.34)	519.51 (31.50)	.89	.17%	.0938	47		
- MLH abroad	1,130.77 (50.62)	1,126.90 (51.84)	-3.87	-.34%	-.2667	48		
- EH abroad	1,649.69 (66.20)	1,633.45 (54.38)	-16.24	-.98%	-.9477	46		
- LH abroad	1,909.90 (54.92)	1,902.37 (54.23)	-7.522	-.39%	-.4873	48		
- ELH abroad	2,232.28 (343.47)	2,179.87 (236.01)	-52.41	-2.35%	-.6288	43		

A6.2 Demonstration B: Changes in the preferences for transport modes at the aggregate level (70-0-30)

Table A6.3 Standard output file for demonstration project B: changes in the preferences for transport modes (70-0-30) – Choice of destination

	Base	Demo B	diff.	%diff	*t*-value	df	S1	S2
* The Netherlands	3,735.30 (60.82)	3,732.92 (55.06)	-2.39	-.06%	-0.2017	32		
♠ metaregion "Water"	1,188.24 (39.98)	1,207.76 (41.78)	19.52	1.64%	2.2121	30	*	
- North	454.77 (22.97)	459.04 (18.40)	4.27	.94%	1.0603	34		
- Mid	187.31 (18.93)	193.04 (19.76)	5.73	3.06%	1.3733	30		
- North Sea	546.16 (28.99)	555.68 (31.91)	9.52	1.74%	1.4201	29		
♠ metaregion "Land north"	1,458.83 (47.20)	1,445.60 (46.44)	-13.23	-.91%	-1.3401	31		
- Mid	549.89 (29.81)	539.80 (29.75)	-10.09	-1.84%	-1.5991	30		
- East	527.16 (32.45)	516.88 (18.27)	-10.29	-1.95%	-2.3837	46	*	
- North	381.76 (25.84)	388.92 (26.04)	7.16	1.87%	1.2963	30		
♠ metaregion "Land south"	877.61 (37.36)	877.04 (35.88)	-.57	-.06%	-0.0745	31		
- West	287.60 (22.97)	293.64 (18.40)	6.04	2.10%	1.2231	30		
- South	590.01 (18.93)	583.40 (19.76)	-6.61	-1.12%	-1.2183	32		
♠ metaregion "Other"	210.63 (20.55)	202.52 (21.60)	-8.11	-3.85%	-1.7785	30		

Table A6.3 Destination choices for demonstration project B- continued

	Base	Demo B	diff.	%diff	*t*-value	df	S1	S2
* Foreign countries	3,392.03 (59.21)	3,393.96 (52.82)	1.93	.06%	.1698	32		
♣ neighbouring countries	1,827.89 (42.82)	1,827.96 (35.75)	.07	.00%	.0084	33		
- France	706.58 (25.59)	702.96 (26.22)	-3.62	-.51%	-.6526	30		
- Belgium & Luxembourg	474.44 (21.32)	477.20 (20.26)	2.77	.58%	.6395	31		
- Germany	445.20 (21.80)	447.96 (23.89)	2.76	.62%	.5487	29		
- United Kingdom	201.68 (15.97)	199.84 (9.89)	-1.84	-.91%	-.8055	42		
♣ more distant countries	1,565.10 (41.69)	1,566.00 (36.97)	.90	.06%	.1131	32		
- Spain & Portugal	432.98 (20.23)	431.88 (18.76)	-1.10	-.25%	-.2728	31		
- Austria & Switzerland	273.07 (15.79)	270.72 (17.68)	-2.36	-.86%	-.6351	29		
- Italy & Greece	282.76 (16.02)	280.24 (16.08)	-2.53	-.89%	-.7407	30		
- Eastern Europe	150.38 (11.74)	150.84 (10.29)	.46	.31%	.2073	32		*
- South East Mediterranean	117.75 (10.34)	122.16 (10.44)	4.41	3.74%	1.9905	30		
- Scandinavia	99.11 (10.15)	100.08 (10.97)	.98	.98%	.4225	29		
- Other	208.08 (13.64)	210.08 (12.57)	2.00	.96%	.7426	32		

Table A6.4 Standard output file for demonstration project B: changes in the preferences for transport modes at the aggregate level (70-0-30) – Transport mode choice

	Base	Demo B	diff.	%diff	*t*-value	df	S1	S2
Means of transport, Av. N° of holidays by								
- car	5,321.90 (63.27)	5,307.92 (56.71)	-13.99	-.26%	-1.1470	32		
- aeroplane	812.95 (29.47)	822.60 (31.87)	9.65	1.19%	1.4391	29		
- alternative, of which	992.48 (32.35)	996.36 (24.45)	3.88	0.39%	.7187	35		
- bus	449.74 (21.04)	450.72 (16.72)	.98	0.22%	.2677	34		
- train	274.26 (16.16)	277.84 (17.49)	3.58	1.31%	.9726	29		
- other	268.48 (16.45)	267.80 (15.28)	-.68	-.25%	-.2080	31		

A6.3 Demonstration C: Ageing of the population

Table A6.5 Standard output file for demonstration project C: Ageing of the population – Composition of the population

	Base	Demo C
Average number of units (per run)	3,562.00 (.00)	3,563.28 (19.09)
HH income		
- <€10,227	307.00 (.00)	330.40 (7.97)
- €10,227-€13,636	389.00 (.00)	422.64 (8.03)
- €13,636-€15,909	416.00 (.00)	417.08 (8.18)
- €15,909-€18,182	531.00 (.00)	528.68 (7.52)
- €18,182-€20,454	431.00 (.00)	420.20 (8.63)
- €20,454-€25,000	587.00 (.00)	578.12 (10.09)
- €25,000-€29,545	367.00 (.00)	352.00 (5.74)
- >€29,545	534.00 (.00)	514.16 (10.63)
Education		
- level 1	282.00 (.00)	270.40 (2.83)
- level 2	729.00 (.00)	752.36 (7.02)
- level 3	363.00 (.00)	372.64 (7.94)
- level 4	535.00 (.00)	538.72 (9.39)
- level 5	227.00 (.00)	225.20 (5.57)
- level 6	803.00 (.00)	784.72 (10.27)
- level 7	476.00 (.00)	475.48 (9.50)
- level 8	147.00 (.00)	143.76 (4.23)
Gender		
- Av. N° of females	1,875.00 (.00)	1,873.52 (14.38)
School holiday region		
- North	1,262.00 (.00)	1,268.68 (13.01)
- Mid	1,016.00 (.00)	1,017.12 (10.84)
- South	1,284.00 (.00)	1,277.48 (10.75)
Province of residence, Av. N° of units from		
- Groningen	142.00 (.00)	144.72 (2.95)
- Friesland	157.00 (.00)	154.76 (5.03)
- Drenthe	138.00 (.00)	142.24 (4.34)
- Overijssel	258.00 (.00)	257.68 (6.00)
- Flevoland	54.00 (.00)	50.88 (2.60)
- Gelderland	410.00 (.00)	405.76 (8.52)
- Utrecht	226.00 (.00)	225.84 (4.54)
- Noord-Holland	506.00 (.00)	511.92 (9.57)
- Zuid-Holland	713.00 (.00)	718.68 (8.92)
- Zeeland	92.00 (.00)	92.48 (2.28)
- Noord-Brabant	585.00 (.00)	576.44 (7.38)
- Limburg	281.00 (.00)	281.88 (5.50)

Table A6.5 Composition of the population for demonstration project C - continued

	Base	Demo C	
Civil state, Av. N° of units			
- married	1,768.00 (.00)	1,810.96 (15.33)	
- cohabiting	218.00 (.00)	198.40 (4.24)	
- divorced	98.00 (.00)	104.68 (4.94)	
- widow(er)	130.00 (.00)	172.64 (5.20)	
- single	317.00 (.00)	308.48 (5.83)	
- other	1,031.00 (.00)	968.12 (7.95)	
Age, Av. N° of			% of population Demo C
- 0-15 year olds	726.00 (.00)	695.28 (5.44)	19.51%
- 16-30 year olds	555.00 (.00)	496.24 (5.80)	13.92%
- 31-55 year olds	1,498.00 (.00)	1,304.44 (11.84)	36.61%
- 56+ year olds	783.00 (.00)	1,067.32 (13.00)	29.95%
HH size, Av. N° of			
- 1-person HH's	495.00 (.00)	537.08 (8.33)	
- 2-person HH's	1,057.00 (.00)	1,178.00 (13.03)	
- 3-person HH's	585.00 (.00)	544.80 (8.16)	
- 4-person HH's	934.00 (.00)	852.52 (9.90)	
- 5+ person HH's	491.00 (.00)	450.88 (6.44)	
Children in HH			
- no children (0-17 years)	1,875.00 (.00)	2,025.36 (17.40)	
- youngest child 6-17 years	1,273.00 (.00)	1,158.80 (12.48)	
- youngest child 0-5 years	414.00 (.00)	379.12 (5.67)	
Av. N° of units with			
- paid jobs	1,580.00 (.00)	1,431.44 (12.36)	
- car in HH	2,922.00 (.00)	2,893.88 (18.38)	
- tourist accommodation with permanent place	255.00 (.00)	263.00 (4.79)	
- tourist accommodation without permanent place	1,279.00 (.00)	1,214.84 (12.07)	
- ski's	338.00 (.00)	331.80 (5.07)	

Table A6.6 Standard output file for demonstration project C: Ageing of the population – Participation choices

	Base	Demo C	diff.	%diff	*t*-value	df	S1	S2
Av. N° of adults (16+ years):	2,836.00 (.00)	2,868.00 (18.65)						
Av. N° of children (0-15 years):	726.00 (.00)	695.28 (5.44)						

Table A6.6 Participation choices for demonstration project C - continued

	Base	Demo C	diff.	%diff	*t*-value	df	S1	S2
For all simulation units, the average N° of (per unit):								
- Day-trips (DT) (adults only)	14.6521 (.163)	14.4823 (.147)	-.170	-1.16%	-5.3799	32	**	**
- Holidays	2.0009 (.021)	1.9860 (.017)	-.015	-0.75%	-4.1221	35	**	**
- Short Breaks (SB)	.7082 (.009)	.6857 (.008)	-.023	-3.18%	-13.1531	33	**	**
- Medium Long Holidays (MLH)	.5694 (.008)	0.5771 (.009)	.008	1.34%	4.0869	30	**	**
- Extended Holidays (EH)	.4038 (.008)	.4021 (.009)	-.002	-0.41%	-.9007	30		*
- Long Holidays (LH)	.2705 (.007)	.2644 (.007)	-.006	-2.25%	-4.1699	30	**	*
- Extra Long Holidays (ELH)	.0489 (.003)	.0566 (.003)	.008	15.66%	12.5274	32	**	**
- days not allocated to tourism (adults)	331.5401 (.368)	331.6068 (.301)	.067	.02%	1.0160	34		
- days not allocated to tourism (children)	348.3092 (.498)	348.5122 (.579)	.203	.06%	1.6773	29		
Av. N° of adults without tourist trips	361.05 (16.19)	406.72 (18.37)	45.670	12.65%	11.8707	29	**	**
Av. N° of children (0-15 years) without tourist trips	113.33 (10.19)	108.64 (12.38)	-4.690	-4.14%	-1.8184	28		
Av. N° of adults making day-trips only	251.68 (15.32)	262.64 (14.86)	10.96	4.35%	3.4653	31	**	**
- Av. N° of day-trips per unit	19.3028 (.527)	19.7896 (.493)	.487	2.52%	4.6191	31	**	**
Av. N° of adults going on holidays only	477.70 (18.36)	486.36 (15.93)	8.655	1.81%	2.5158	33	*	
- Av. N° of holidays per unit	2.4091 (.082)	2.4010 (.072)	-.008	-.33%	-.5216	32		
- Av. N° of SB per unit	.8127 (.060)	.7796 (.053)	-.033	-4.07%	-2.9145	32	**	
- Av. N° of MLH per unit	.6679 (.025)	.6799 (.023)	.012	1.78%	2.4670	32	*	*
- Av. N° of EH per unit	.4884 (.024)	.4860 (.018)	-.002	-.49%	-.5857	36		
- Av. N° of LH per unit	.3541 (.022)	.3628 (.023)	.009	2.47%	1.8328	30		**
- Av. N° of ELH per unit	.0860 (.012)	.0928 (.011)	.007	7.84%	2.8525	32	**	
- Av. N° of days not allocated to tourism per unit	341.6551 (.754)	341.3044 (.790)	-.351	-.10%	-2.1018	30	*	**

Table A6.6 Participation choices for demonstration project C- Continued

	Base	Demo C	diff.	%diff	*t*-value	df	S1	S2
Av. N° of children (0-15 years) going on holidays	612.67 (10.19)	586.64 (12.52)	-26.03	-4.25%	-9.9899	28	**	**
- Av. N° of holidays per unit	2.1387 (.039)	2.1284 (.042)	-.010	-0.49%	-1.1829	29		
- Av. N° of SB per unit	.7457 (.020)	.7492 (.022)	.004	0.48%	.7779	29		
- Av. N° of MLH per unit	.5985 (.018)	.5963 (.026)	-.002	-0.36%	-.3931	27		
- Av. N° of EH per unit	.4355 (.021)	.4327 (.018)	-.003	-0.65%	-.7207	32		
- Av. N° of LH per unit	.3308 (.017)	.3221 (.022)	-.009	-2.63%	-1.9143	28		
- Av. N° of ELH per unit	.0283 (.007)	.0280 (.005)	.000	-1.04%	-.2706	37		
- Av. N° of days not allocated to tourism per unit	345.2209 (.524)	345.4593 (.545)	.238	0.07%	2.0708	30	*	
Av. N° of adults going on holidays & day-trips	1,745.57 (21.52)	1,712.28 (22.92)	-33.285	-1.91%	-6.8925	30	**	**
- Av. N° of day-trips per unit	21.0224 (.090)	21.2221 (.099)	.200	.95%	9.6452	29	**	**
- Av. N° of holidays per unit	2.6732 (.029)	2.7216 (.030)	.048	1.81%	7.6207	30	**	**
- Av. N° of SB per unit	.9611 (.016)	.9489 (.016)	-.012	-1.28%	-3.6226	30	**	**
- Av. N° of MLH per unit	.7691 (.012)	.8036 (.011)	.034	4.48%	14.4291	31	**	**
- Av. N° of EH per unit	.5375 (.011)	.5505 (.012)	.013	2.42%	4.9645	29	**	**
- Av. N° of LH per unit	.3391 (.011)	0.3368 (.008)	-.002	-.67%	-1.2691	36		
- Av. N° of ELH per unit	.0664 (.005)	.0819 (.005)	.015	23.34%	13.8272	30	**	**
- Av. N° of days not allocated to tourism per unit	319.8098 (.388)	318.8350 (.378)	-.975	-.30%	-12.1276	31	**	**

Table A6.7 Standard output file for demonstration project C: Ageing of the population – Holiday choices

	Base	Demo C	diff.	%diff	*t*-value	df	S1	S2
Av. duration of holidays (in days)	9.18 (.06)	9.28 (.06)	.100	1.09%	7.4295	31	**	**

Table A6.7 Holiday choices for demonstration project C- Continued

	Base	Demo C	diff.	%diff	*t*-value	df	S1	S2
Travel party, Av. N° of holidays with								
- alone	288.26	291.16	2.90	1.00%	.6768	29		
	(17.91)	(20.43)						
- adults only	2,756.85	2,879.00	122.15	4.43%	11.0625	29	**	**
	(47.18)	(52.63)						
- schoolchildren	1,425.93	1,314.72	-111.21	-7.80%	-16.1782	32	**	**
(6-14 years)	(36.26)	(31.89)						
- other children	981.63	906.96	-74.68	-7.61%	-13.5275	32	**	**
(0-5 or 15+ years)	(28.88)	(25.64)						
- party of 9+ people	750.48	745.84	-4.64	-.62%	-.8697	32		**
	(27.26)	(24.84)						
- unknown	924.18	938.76	14.58	1.58%	2.4457	31	*	*
	(29.66)	(27.90)						
Season & In-/outside school holiday period, Av. N° of holidays in								
- winter, during	435.61	428.52	-7.09	-1.63%	-1.9318	34		
school holiday period	(20.73)	(16.82)						
- winter, outside	464.72	480.28	15.56	3.35%	3.0033	30	**	**
school holiday period	(24.25)	(24.44)						
- spring, during	1,028.88	999.52	-29.36	-2.85%	-5.4904	35	**	*
school holiday period	(31.04)	(24.38)						
- spring, outside	1,563.60	1,588.88	25.28	1.62%	3.1294	31	**	**
school holiday period	(38.58)	(38.02)						
- summer-begin	832.73	806.16	-26.57	-3.19%	-4.6306	29	**	**
	(23.57)	(27.45)						
- summer-mid	1,059.38	1,021.48	-37.90	-3.58%	-5.3649	30	**	**
	(31.54)	(33.52)						
- summer-end	637.14	626.92	-10.23	-1.60%	-1.7830	29		
	(25.35)	(27.24)						
- autumn, during	334.27	329.24	-5.04	-1.51%	-1.6516	38		**
school holiday period	(19.68)	(13.56)						
- autumn, outside	771.00	795.44	24.44	3.17%	4.9382	33	**	*
school holiday period	(27.25)	(22.79)						
Destination, Av. N° of holidays to/in:								
- The Netherlands	3,735.30	3,719.80	-15.51	-.42%	-1.2334	31		
	(60.82)	(59.06)						
- metaregion "Water"	1,188.24	1,174.24	-14.01	-1.18%	-1.7001	31		**
	(39.98)	(38.69)						
- North	454.77	452.00	-2.78	-.61%	-.6162	32		
	(22.97)	(21.00)						
- Mid	187.31	182.68	-4.63	-2.47%	-1.1874	31		*
	(18.93)	(18.31)						
- North Sea	546.16	539.56	-6.60	-1.21%	-1.0309	30		
	(28.99)	(30.33)						

Table A6.7 Holiday choices for demonstration project C- Continued

	Base	Demo C	diff.	%diff	*t*-value	df	S1	S2
- metaregion "Land north"	1,458.83 (47.20)	1,474.92 (55.09)	16.10	1.10%	1.3981	29		
- Mid	549.89 (29.81)	553.16 (25.78)	3.27	.59%	.5862	33		
- East	527.16 (32.45)	528.48 (35.10)	1.32	.25%	.1781	29		*
- North	381.76 (25.84)	393.28 (29.61)	11.52	3.02%	1.8580	29		
- metaregion "Land south"	877.61 (37.36)	858.88 (31.91)	-18.73	-2.13%	-2.7117	33	*	
- West	287.60 (22.97)	281.88 (21.00)	-5.72	-1.99%	-1.1387	30		
- South	590.01 (18.93)	577.00 (18.31)	-13.01	-2.21%	-2.0387	30		
- metaregion "Other"	210.63 (20.55)	211.76 (19.07)	1.14	.54%	.2781	31		
* Foreign countries	3,392.03 (59.21)	3,356.64 (60.03)	-35.39	-1.04%	-2.7834	30	**	
♣ neighbouring countries	1,827.89 (42.82)	1,792.92 (52.47)	-34.98	-1.91%	-3.2025	28	**	**
- France	706.58 (25.59)	689.04 (22.25)	-17.54	-2.48%	-3.6516	32	**	**
- Belgium & Luxembourg	474.44 (21.32)	455.48 (30.54)	-18.96	-4.00%	-3.0133	27	**	**
- Germany	445.20 (21.80)	449.44 (25.95)	4.24	.95%	.7823	28		
- United Kingdom	201.68 (15.97)	198.96 (13.73)	-2.72	-1.35%	-.9142	33		
♣ more distant countries	1,565.10 (41.69)	1,563.72 (34.69)	-1.38	-.09%	-.1831	33		
- Spain & Portugal	432.98 (20.23)	426.88 (21.99)	-6.10	-1.41%	-1.3182	29		
- Austria & Switzerland	273.07 (15.79)	270.76 (20.89)	-2.32	-.85%	-.5352	28		
- Italy & Greece	282.76 (16.02)	282.24 (16.32)	-0.53	-.19%	-.1519	30		
- Eastern Europe	150.38 (11.74)	148.00 (11.24)	-2.38	-1.58%	-.9934	31		
- South East Mediterranean	117.75 (10.34)	125.28 (9.39)	7.53	6.39%	3.7345	32	**	
- Scandinavia	99.11 (10.15)	101.56 (9.15)	2.46	2.48%	1.2487	32		
- Other	208.08 (13.64)	209.00 (12.54)	.92	.44%	.3423	32		

Table A6.7 Holiday choices for demonstration project C- Continued

	Base	Demo C	diff.	%diff	*t*-value	Df	S1	S2
Accommodation, Av. N° of holidays in								
- hotel	2,440.68	2,458.04	17.36	.71%	1.6583	32		
	(53.30)	(48.83)						
- non-permanent-own	921.24	895.12	-26.12	-2.84%	-4.3646	32	**	**
	(30.87)	(27.86)						
- permanent-own	975.52	1,004.56	29.04	2.98%	3.8528	31	**	**
	(37.15)	(35.32)						
- other	2,789.89	2,718.72	-71.18	-2.55%	-6.8603	32	**	**
	(54.92)	(48.10)						
Means of transport, Av. N° of holidays by								
- car	5,321.90	5,263.20	-58.71	-1.10%	-4.9290	32	**	**
	(63.27)	(55.19)						
- aeroplane	812.95	816.84	3.89	.48%	.6905	32		
	(29.47)	(26.17)						
- alternative, of which	992.48	996.40	3.92	.39%	.5720	30		
	(32.35)	(32.30)						
- bus	449.74	452.28	2.54	.56%	.6884	34		
	(21.04)	(16.88)						
- train	274.26	278.96	4.70	1.71%	1.2636	29		**
	(16.16)	(17.70)						
- other	268.48	265.16	-3.32	-1.24%	-.8102	28		
	(16.45)	(19.65)						
Day of departure, Av. N° of holidays departing on								
- Friday/Saturday	3,922.80	3,855.48	-67.33	-1.72%	-6.0159	33	**	**
	(61.61)	(51.54)						
- other day of the week	3,204.53	3,220.96	16.43	.51%	1.3965	30		
	(52.75)	(55.79)						
Average expenditures (NLG) per								
* holiday	744.50	753.97	9.47	1.27%	4.2162	32	**	**
	(11.51)	(10.46)						
♠ domestic holiday	289.04	291.06	2.02	.70%	1.7766	29		
	(4.75)	(5.44)						
- domestic SB	180.01	180.48	.47	.26%	.5539	30		
	(3.87)	(4.04)						
- domestic MLH	321.02	321.12	.09	.03%	.0543	31		
	(8.35)	(8.02)						
- domestic EH	491.47	489.12	-2.36	-.48%	-.4977	30		
	(21.47)	(22.41)						
- domestic LH	559.49	548.81	-10.68	-1.91%	-1.2352	28		**
	(33.17)	(41.60)						
- domestic ELH	541.43	547.23	5.80	1.07%	.3780	31		
	(75.17)	(71.97)						

Table A6.7 Holiday choices for demonstration project C- Continued

	Base	Demo C	diff.	%diff	*t*-value	df	S1	S2
♣ holiday abroad	1,246.08 (20.40)	1,267.15 (23.70)	21.065	1.69%	4.2511	29	**	**
- SB abroad	507.23 (25.39)	499.36 (28.17)	-7.878	-1.55%	-1.3325	29		
- MLH abroad	1,038.07 (35.43)	1,059.73 (38.51)	21.652	2.09%	2.6732	29	*	
- EH abroad	1,503.14 (40.25)	1,515.19 (35.51)	12.051	0.80%	1.5751	32		
- LH abroad	1,685.43 (45.97)	1,711.81 (32.30)	26.372	1.56%	3.6467	37	**	**
- ELH abroad	2,043.39 (106.61)	2,063.27 (96.23)	19.888	0.97%	0.9622	32		

Subject Index

D

E

G

H

I

L

M

N

P

Q

R

S

T

U

V

W

Author Index

MERLIN: Een Beslissingsondersteunend Systeem voor Toeristisch-Recreatieve Planning

(Ontwikkeling en test van een regelgebaseerd microsimulatie model voor de evaluatie van alternatieve ontwikkelingsscenarios en planningsmogelijkheden)

Ter ondersteuning van ruimtelijke planning en beleid voor recreatie en toerisme hebben vele studies de voorkeuren van (potentiële) bezoekers voor bestemmingen onderzocht. Voor het kwantificeren van deze voorkeuren, en vooral voor het voorspellen van toekomstig keuzegedrag wordt vaak gebruik gemaakt van zogeheten *revealed* en/of *stated preference* en *choice* modellen. Op basis van de geobserveerde voorkeuren voor werkelijke of hypothetische bestemmingen worden in deze modellen onder de aanname van nutsmaximaliserend gedrag de waarderingen van de verschillende kenmerken van bestemmingen vastgesteld. Met behulp van deze waarderingen kunnen vervolgens uitspraken gedaan worden over de voorkeuren voor nieuwe bestemmingen.

In de jaren '80 en '90 is er veel kritiek geuit op deze benadering van keuzegedrag in de vrije tijd. Allereerst werd duidelijk dat het toeristisch[47] keuzeproces bestaat uit meer dan alleen de bestemmingskeuze. Zo neemt de bezoeker ook beslissingen over hoe vaak hij of zij toeristische uitstapjes maakt, met wie hij of zij zal reizen, met welk vervoermiddel, op welk tijdstip, en ga zo maar door. Belangrijker is het inzicht dat al deze `deelkeuzes´ of facetten van het keuzeproces elkaar sterk kunnen beïnvloeden. Zo zal voor een dagtocht met kinderen vaak gezocht worden naar een kindvriendelijke bestemming en zal een reis naar Australië niet licht op de fiets ondernomen worden. Ook werd duidelijk dat er interacties bestaan tussen de facetten van verschillende uitstapjes. De keuze voor een lange, verre reis in het voorjaar, kan bijvoorbeeld resulteren in de keuze voor een goedkopere en/of kortere vakantie in het najaar: je kunt je inkomen en vrije dagen tenslotte maar één keer besteden. Tot slot bleek de aanname van nuts-maximaliserend gedrag in deze complexe keuzesituaties moeilijk te verdedigen.

Het *doel* van deze studie is daarom het ontwikkelen en testen van een model dat rekening houdt met de bovengenoemde complexiteit van toeristisch keuzegedrag. Het te ontwikkelen instrument moet diverse toekomstscenarios door kunnen rekenen op de gevolgen voor toeristische uitstapjes in al haar facetten. Er

[47] In dit proefschrift wordt, mede ter bevordering van de leesbaarheid, steeds gesproken over *toeristische* uitstapjes waar het om zowel dagtochten als om vakanties gaat. Een deel van deze uitstapjes zou echter beter aangeduid kunnen worden als (toeristisch-)recreatieve uitstapjes.

is gekozen voor *microsimulatie* omdat dit de mogelijkheid biedt om complexere modellen te bouwen door het combineren van hypotheses en modellen die de onderdelen van het (keuze)gedrag beschrijven. Hierbij wordt elke fase in het keuzeproces van een individu beschreven aan de hand van een verdeling of een model. Effecten op aggregaat niveau worden verkregen door het combineren van de individuele simulatieresultaten. Het ontwikkelde model is onderdeel van het zogenoemde *MERLIN*-systeem dat met een menu-gestuurde user-interface de mogelijkheid biedt om onder verschillende condities jaarlijkse patronen van toeristische uitstapjes te simuleren en te vergelijken met de huidige situatie.

Om dit doel te bereiken moeten allereerst drie *onderzoeksvragen* beantwoord worden. Deze onderzoeksvragen lopen als een rode draad door de theoretische hoofdstukken 2 tot en met 5. Allereerst dient er een conceptueel model ontwikkeld te worden van het gedrag dat gesimuleerd wordt. Het gaat hierbij om vragen zoals: Hoe worden keuzes voor toeristische uitstapjes gemaakt? Welke fases in keuze-proces spelen een rol? En hoe beïnvloeden keuzes in verschillende fases elkaar? Dit is het onderwerp van *hoofdstuk 2.* Allereerst wordt een algemeen model van consumenten-keuzegedrag besproken. Daarna wordt het meer traditionele onderzoek naar toeristische bestemmingskeuzes uitgewerkt. Tot slot wordt ingegaan op de meer recente inzichten in en waardering van de complexiteit van het toeristisch keuzeproces. Een vrij algemeen geaccepteerde voorstelling van dit proces is dat het bestaat uit een groot aantal samenhangende beslissingen over de diverse facetten van toeristische uitstapjes. Er wordt ook wel gesproken over het *profiel* van een toeristisch uitstapje. Dit proces wordt geconditioneerd door een complex samenspel van persoons- en huishoudvoorkeuren (preferenties), beperkingen (restricties), kansen (mogelijkheden) en eerdere beslissingen. Ook wordt verondersteld dat verschillende uitstapjes elkaar kunnen beïnvloeden. Dit laatste betekent dat uitstapjes van een individu over een langere periode en in samenhang met elkaar bestudeerd moeten worden. Op basis van deze complexe voorstelling van het toeristisch keuzeproces presenteert hoofdstuk 2 een conceptueel model voor jaarlijkse patronen van toeristische uitstapjes. Het gaat hierbij zowel om dagtochten als om vakanties.

Een tweede onderzoeksvraag die beantwoord moet worden is hoe elk facet van het toeristisch keuzeproces het best beschreven kan worden met een model. Hierbij moet ook aandacht worden besteed aan de vraag hoe relaties tussen verschillende facetten en opeenvolgende toeristische uitstapjes beschreven kunnen worden. Hiertoe verkent *hoofdstuk 3* de model-typen die in de afgelopen decennia zijn gebruikt om toeristisch keuzegedrag te analyseren en te beschrijven op basis van keuzetheorieën. Er wordt onderscheid gemaakt tussen nutsgebaseerde en niet-nutsgebaseerde modellen. *Nutsgebaseerde modellen* gaan ervan uit dat mensen het nut dat zij ontlenen aan een bepaald keuze-alternatief willen maximaliseren. Het meest bekende nutsgebaseerde model voor discrete keuzes is het Multinomial Logit (of MNL-) model. Aanhangers van *niet-nutsgebaseerde modellen*, daarentegen, denken dat mensen niet in staat zijn om nutsmaximaliserend gedrag te vertonen omdat (1) mensen door diverse beperkingen niet altijd dat alternatief kunnen kiezen dat het hoogste nut oplevert (restrictie-gebaseerde benaderingen); (2) de werking van het menselijk brein anders werkt (Neurale Netwerken); en/of (3) aangenomen wordt dat mensen gebruik maken van relatief eenvoudige regels die aangeven welke keuzes in het verleden onder vergelijkbare omstandigheden redelijke oplossingen opleverden (regelgebaseerde modellen). Op basis van een verkenning van de verschillende model-typen geeft hoofdstuk 3 tot slot aan hoe elk keuzefacet gemodelleerd zal worden (zie hiervoor de hoofdstukken 7 tot en met 11).

De derde en laatste onderzoeksvraag heeft betrekking op de structuur van het simulatiemodel: hoe kunnen de verschillende modellen gecombineerd worden tot één (simulatie)model? Hiertoe wordt in *hoofdstuk 4* allereerst een overzicht gegeven van de voor- en nadelen van het gebruik van simulatie. Er worden ook diverse indelingen gegeven waarmee simulatiemodellen beschreven kunnen worden. Simulatie wordt hierbij opgevat als een klasse van modellen voor het onderzoeken van systemen die te complex zijn voor analytische beschrijvingen. Uitgaande van deze definitie geeft hoofdstuk 4 een overzicht van bestaande simulatiemodellen op het gebied van recreatie en toerisme. Hierbij worden drie buitenlandse en twee Nederlandse voorbeelden geïdentificeerd en beschreven. Tot slot wordt *MERLIN* gekarakteriseerd als een stochastisch, discreet, statisch en empirisch getest microsimulatie model dat geïmplementeerd wordt met behulp van een algemene programmeertaal.

Hoofdstuk 5 voltooit de theoretische ontwikkeling van *MERLIN* met de formalisatie van het te modelleren probleem en een beschrijving van de architectuur van het systeem. De architectuur van het *MERLIN*-systeem bestaat uit vier hoofdcomponenten. Allereerst is er een component die het toeristisch keuzegedrag van een gegeven populatie simuleert onder verschillende condities en gebaseerd op een gegeven set van (nuts- en regelgebaseerde) modellen. Deze component zorgt er ook voor dat de resultaten van dit simulatieproces worden weggeschreven voor diverse toepassingen. Een tweede component genereert op basis van een door de gebruiker geformuleerd project de populatie (en de systeem en institutionele condities) waarvan het toeristisch gedrag gesimuleerd wordt. Een derde component heeft als taak om op basis van de simulatie-resultaten kruistabellen te genereren. De laatste component is de user-interface. Deze component zorgt ervoor dat de gebruiker de condities kan bepalen waaronder het keuzegedrag van een bepaalde populatie wordt gesimuleerd. Ook het inspecteren van de simulatie-resultaten loopt voornamelijk via de user-interface.

Hoofdstuk 6 beschrijft de dataverzameling en vormt daarmee de brug tussen de theoretische en de daaropvolgende empirische hoofdstukken. Het hoofdstuk begint met de beschrijving van de benodigde data en een pilot-studie om het verzamelen ervan te testen. Daarna volgt een beschrijving van de echte data-verzameling. De belangrijkste bron is het Continu Vakantie Onderzoek (CVO) uit 1998, aangevuld met extra vragen over dagtochten en (her-) planningsgedrag. Het CVO is een uitgebreid onderzoek dat sinds 1980 jaarlijks het vakantiegedrag van de Nederlandse bevolking registreert. Dit gebeurt door vier keer per jaar het vakantiegedrag van een representatief panel te registeren. In 1998 zijn in de eerste meeting 5.151 personen benaderd. Na vier metingen en vier maal non-response bestond de netto steekproef uit 3562 personen. Omdat het CVO geen informatie verzamelt over dagtochten zijn in de 4^e^ kwartaalmeting (in december 1998) extra vragen toegevoegd over het aantal dagtochten en vakanties van de respondenten in 1998. Ook is er informatie verzameld over hoe toeristische keuzes bijgesteld worden onder invloed van veranderende omstandigheden (re-scheduling of -allocatiegedrag). Het ging hierbij om veranderingen zoals andere werktijden, meer of minder vrije dagen, inkomensveranderingen en veranderingen in het aanbod aan toeristisch-recreatieve faciliteiten of hulpmiddelen. Deze extra vragen hebben alleen betrekking op (veranderingen in) het jaarlijks aantal toeristisch-recreatieve uitstapjes en de duur daarvan. Tot slot wordt een overzicht gegeven van de verzamelde data. Hieruit blijkt onder andere dat er in sommige gevallen afwijkingen bestaan tussen de gegevens over het aantal toeristische uitstapjes per jaar (uitgesplitst naar duur) in de bruikbare antwoorden op de extra vragen en die voor de hele populatie (op basis van de CVO gegevens) over 1998. In hoofdstuk 12 worden gewichten geïntroduceerd om deze afwijkingen te corrigeren.

Na de beschrijving van de data volgen vijf empirische hoofdstukken die steeds voor de opeenvolgende keuzefacetten één of meerdere empirische modellen afleiden en beschrijven. Het eerste empirische hoofdstuk, *hoofdstuk 7*, beschrijft een algemeen model voor de allocatie van tijd aan toeristische uitstapjes. Hiermee kan MERLIN het jaarlijkse programma van toeristische uitstapjes, inclusief de duur ervan, genereren. Dit proces bestaat uit twee fasen. Allereerst besluit de (potentiële) reiziger of hij of zij in een bepaald jaar dagtochten en/of vakanties wil maken (participatiekeuze). Op basis van een analyse van de relatie tussen dagtochten, korte vakanties (2-4 dagen) en langere vakanties (5+ dagen) in de pilot-studie is hierbij gekozen voor de keuze tussen (1) alleen dagtochten; (2) alleen vakanties; (3) zowel dagtochten als vakanties; of (4) geen toeristische uitstapjes. Indien de reiziger kiest voor één van de eerste drie alternatieven, beschrijft een model voor uitstapjeskeuzes vervolgens de allocatie van tijd aan dagtochten en/of vakanties van verschillende duur (2-4 dagen; 5-8 dagen; 9-15 dagen; 16-28 dagen; en 29 of meer dagen). De verklarende variabelen in de modellen bestaan uit (1) het basisnut van elk alternatief; (2) socio-economische en –demografische kenmerken van de (potentiële) reiziger en zijn of haar huishouden; en (3) de toestand van de besluitvormingscontext (een dummy die aangeeft of de gegevens betrekking hebben op de uitstapjes in 1998 of op het re-allocatiegedrag; zie hoofdstuk 6).

In hoofdstuk 3 is beargumenteerd dat dit getrapte keuzeproces het best beschreven kan worden met een hierarchisch logit model, met name het nested-logit model. Het bleek echter niet mogelijk dit model te schatten. De waarde van de parameter van de verbindende variabele lag namelijk buiten het vereiste bereik. Er is daarom een set van sequentiële MNL-modellen geschat. Er is zowel een set voor volwassenen als één voor kinderen geschat omdat er voor kinderen geen gegevens beschikbaar waren over dagtochten en re-allocatiegedrag. Inkomen, opleiding, leeftijd en het bezit van toeristische accommodaties (met of zonder vaste standplaats) bleken de belangrijkste verklarende variabelen. Ook de dummies voor de re-scheduling of –allocatiescenario's bleken belangrijke effecten te hebben.

Gegeven het aantal en de duur van de uitstapjes worden beslissingen genomen over de keuzefacetten. Voor het beschrijven van deze keuzes wordt gebruikt gemaakt van probabilistische regelgebaseerde modellen die vastgelegd zijn in zogenoemde beslistabellen. Een beslistabel geeft aan onder welke condities (ALS <condities>) welke keuze (DAN <keuze>) met een bepaalde mate van waarschijnlijkheid gemaakt zal worden. De relaties tussen de verschillende keuzefacetten worden hierbij opgenomen door keuzes in eerdere fases op te nemen als condities voor latere keuzes. Er wordt een vaste volgorde van de keuzefacetten aangenomen. Voor het afleiden van regels uit empirische gegevens wordt gebruik gemaakt van een algoritme gebaseerd op de klassieke Chi-kwadraat Automatische Interactie Detectie (CHAID-)analyse. Dit algoritme is een opvolger van het AID-algoritme en werd rond 1980 door Kass ontwikkeld voor nominale en ordinale variabelen. Toegepast op de inductie van beslisregels splitst dit algoritme een set van observaties steeds in beslisregels die homogener zijn ten aanzien van de keuze-alternatieven. Dit gebeurt door de significantie van de Chi-kwadraat (χ^2) te maximaliseren. In dit proces wordt de *p*-waarde van de χ^2 vermenigvuldigd met de zogeheten Bonferonni-multiplier omdat, ter minimalisatie van het aantal regels, de categorieën van elke conditie-variabelen in een eerdere stap zo optimaal mogelijk bij elkaar zijn gevoegd. Dit proces wordt herhaald voor elke beslisregel tot één van de vooraf geformuleerde stopcriteria bereikt is. Op deze wijze creëert het algoritme een complete set van wederzijds uitsluitende en uitputtende beslisregels. Elke regel wordt hierbij gedefinieerd door een unieke combinatie van condities (d.w.z. categorieën van de condititie-variabelen) en een bepaalde (kans-)verdeling over de keuze-alternatieven. Een beslisregel wordt in een beslistabel weergegeven als één kolom. In het simulatieproces wordt vervolgens met behulp van het trekken van aselecte getallen bepaald welk keuze-alternatief uiteindelijk gekozen wordt (Monte Carlo simulatie).

Hoofdstuk 8, het tweede empirische hoofdstuk, gaat over de keuze van het reisgezelschap. Aangezien er geen details beschikbaar waren over de dagtochten, wordt dit keuzefacet (en alle volgende) alleen beschreven voor vakanties. Het afleiden van beslisregels voor reisgezelschappen (en alle volgende keuzefacetten) verloopt in stappen. Als eerste stap wordt de keuzeset voor reisgezelschappen gedefinieerd. Deze keuzeset is gebaseerd op de aanwezigheid en de leeftijd van kinderen en op de omvang van het gezelschap. Het gaat om de volgende opties: (1) alleen; (2) alleen volwassenen (20 jaar en ouder); (3) met schoolgaande kinderen (6-14 jaar); (4) met andere kinderen (0-5 of 15-19 jaar); (5) gezelschappen van 9 of meer personen (ongeacht de leeftijd); en (99) onbekend. Dit laatste alternatief komt voort uit het feit dat de CVO-meting alleen alle details registreert van de twee langste vakanties per kwartaal. Indien een respondent dus meer dan 2 vakanties maakt in een kwartaal, ontbreekt van de 3e en de volgende vakantie(s) een aantal gegevens, waaronder het reisgezelschap.

De tweede stap is het definiëren van de conditievariabelen op basis waarvan de beslisregels afgeleid worden. In principe zijn er vier typen conditievariabelen voor alle keuzefacetten: (1) persoons- en huishoudkenmerken van de toerist; (2) de keuzes die de toerist al gemaakt heeft over het aantal dagtochten en vakanties in dat jaar; (3) de keuzes die al gemaakt zijn over andere keuzefacetten van de betreffende vakantie; en (4) de keuzes die al gemaakt zijn over andere vakanties. Voor de keuze van het reisgezelschap is er pas één andere keuze gemaakt over de betreffende vakantie, namelijk de lengte van de vakantie in dagen. Er wordt nog geen rekening gehouden met de keuzes die al gemaakt zijn over andere vakanties.

De derde stap stelt de stopcriteria voor het op CHAID gebaseerde algoritme vast. Het eerste criterium heeft betrekking op het minimale significantie-niveau van de χ^2 dat vereist is om een beslisregel verder op te splitsen. Voor alle beslistabellen is dit 5%. Het tweede criterium heeft betrekking op het aantal observaties dat minimaal aanwezig moet zijn voor en na de splitsing. Hiervoor wordt steeds een gevoeligheidsanalyse uitgevoerd waarbij gelet wordt op de prestaties van diverse modellen als het gaat om het terugvoorspellen van de oorspronkelijke observaties. Voor de keuze van het reisgezelschap, bijvoorbeeld, is gekozen voor minimaal 100 observaties voor en 45 observaties na het opsplitsen van een beslisregel.

In de vierde stap worden de beslisregels afgeleid. Deze beslisregels worden weergegeven in een beslistabel, geïnterpreteerd en (in grote lijnen) besproken. Voor de keuze van het reisgezelschap, bijvoorbeeld, bleek de aanwezigheid van kinderen in het huishouden en het bezit van een vaste toeristische accommodatie belangrijk. Tot slot worden voor elke beslistabel enkele belangrijke maten voor de validatie besproken. Op basis van deze maten kan voor de 68 beslisregels voor het reisgezelschap geconcludeerd worden dat het model een belangrijke verbetering betekend voor de mate waarin gezelschapskeuzes voorspeld kunnen worden.

De hoofstukken 9, 10 en 11 behandelen op dezelfde wijze beslistabellen voor de timing, de bestemming, de logiesaccommodatie en de uitgaven. Nieuw daarbij is dat ook conditie-variabelen opgenomen worden die iets zeggen over de keuzes die al genomen zijn voor andere vakanties over het betreffende keuzefacet. Op basis van de duur en het reisgezelschap wordt namelijk de planningsprioriteit van een vakantie binnen het programma van toeristische uitstapjes bepaald. De reden hiervoor is dat beslissingen voor langere vakanties en/of vakanties met mensen die beperkt zijn in hun keuzevrijheid eerder genomen zullen worden en dus de keuzes beïnvloeden die volgen. Door het opnemen van sommatievariabelen kan herhalings- en variatiezoekend gedrag afgeleid worden (als die aanwezig zijn).

De keuze-alternatieven voor de timing van vakanties worden in *hoofdstuk 9* dusdanig geoperationaliseerd dat zij relevant zijn voor de piekperioden waarmee de sector te maken heeft. Het gaat hierbij om (1) de keuze om tijdens een schoolvakantie-periode op vakantie te gaan of daarbuiten; (2) de seizoenskeuze; en (3) de keuze van de dag van vertrek (vrijdag of zaterdag, of op een rustigere dag van de week). De keuze voor de schoolvakantieperiode wordt in belangrijke mate bepaald door de aanwezigheid van kinderen in het huishouden, de planningsprioriteit van de vakantie en het reisgezelschap. Voor de keuze van het seizoen, bijvoorbeeld, zijn de belangrijkste conditionerende variabelen de keuze om in een schoolvakantieperiode te reizen, het reisgezelschap en het bezit van ski's. Daarnaast zijn veel sommatievariabelen in de beslisregels opgenomen omdat veel mensen ervoor kiezen om vakanties over de seizoenen te spreiden. De keuze van de vertrekdag, tenslotte, wordt pas gemaakt als alle andere keuzefacetten (behalve de uitgaven) overwogen zijn. De belangrijkste condities hierbij zijn de bestemming (binnenland of buitenland), het aantal vakanties dat al of vrijdag of zaterdag is begonnen (veelal leidend tot herhalingsgedrag), de duur van de vakantie en het seizoen waarin gereisd wordt.

Hoofdstuk 10 gaat over beslisregels voor bestemmingskeuzen. Deze keuzes worden in MERCIN voorgesteld als een gefaseerd proces waarin de toerist systematisch bepaalde (groepen van) bestemmingen uitsluit op basis van zijn of haar persoons- en huishoudvoorkeuren en -beperkingen in relatie tot de eigenschappen van de (potentiële) bestemmingen. Hierbij wordt allereerst de keuze gemaakt voor een buiten- of binnenlandse vakantie. Gegeven de keuze om naar het buitenland te gaan, kiest de toerist vervolgens voor een buurland (Frankrijk, België/Luxemburg, Duitsland en het Verenigd Koninkrijk/Ierland) of een bestemming die geografisch of cultureel verder van Nederland af ligt. Vervolgens wordt de keuze gemaakt voor een specifiek buurland of een bepaalde groep van verderweg gelegen landen (bijvoorbeeld Oostenrijk/ Zwitserland, Scandinavië/Denemarken, het voormalig Oost-blok, etc.). Bij de keuze voor een bestemming binnen Nederland, wordt allereerst de keuze gemaakt tussen vier zogenoemde meta-regio's: water (Noordzeekust, watergebieden in het noorden en watergebieden in het midden van Nederland); noordelijke landgebieden (Utrecht, 't Gooi en de Veluwe, de Achterhoek, Twente, Salland en de Vechtstreek en de zandgebieden van Groningen, Friesland en Drenthe); zuidelijke landgebieden (west en midden-Brabant en oost-Brabant, het Rijk van Nijmegen en Limburg); en overige gebieden (de vier grote steden en overig Nederland). Pas daarna volgt de keuze voor een specifiek toeristengebied. Elk van de genoemde keuzes in dit proces wordt beschreven door een beslistabel. Bestemmingskeuzes worden aldus beschreven door een hiërarchische set van acht beslistabellen.

De conditievariabelen voor de bestemmingskeuzes omvatten naast de gebruikelijke typen ook een aantal condities die het mogelijk maken om de beslisregels voor bestemmingskeuzes "vooruit te laten kijken" naar keuzefacetten die later in het keuzeproces beschouwd worden. Allereerst is een variabele opgenomen die aangeeft hoeveel geld er reeds is uitgegeven aan (of eigenlijk: gepland voor) vakanties met een hogere planningsprioriteit. Hierdoor is het mogelijk om eventuele financiële relaties tussen vakantiebestemmingen in planningsproces te beschrijven. De tweede manier heeft betrekking op de relatie tussen bestemmings- en vervoermiddelkeuzes. De gedachte hierachter is dat eventuele veranderingen in de voorkeuren voor vervoermiddelen (bijvoorbeeld door veranderingen in de kosten of het belang van milieu-overwegingen in het keuzeproces) grote invloed kunnen hebben op de bestemmingskeuzes. Hiertoe zijn variabelen gevormd die de algemene (voorspelde) geneigdheid van een toerist weergeven om een bepaald vervoermiddel te kiezen (dus onafhankelijk van de reeds bekende kenmerken van de vakantie). Elke "geneigdheidsvariabele" geeft op een schaal van 0 tot 100 de voorkeur weer voor de auto, het vliegtuig en een alternatieve vorm van vervoer; samen tellen de drie “geneigdheidsvariabelen” voor één individu altijd op tot 100. De waarden van drie variabelen worden voorspeld met behulp van een MNL-model. Dit model is gekalibreerd door de in 1998 geobserveerde geneigdheden voor de vervoerswijzen (in de CVO data) te relateren aan (1) aggregate voorkeuren voor de drie vervoerswijzen (in 1998: 74,9% auto, 12,2% vliegtuig en 12,9% alternatief); (2) diverse persoons- en huishoud-kenmerken; en (3) het aantal vakanties. Indien de “geneigdheidsvariabelen” geselecteerd worden in de beslisregels kan de gebruiker scenarios simuleren waarin de aggregate voorkeuren voor de vervoerswijzen veranderen t.o.v. 1998.

Uit de bespreking van de beslistabellen voor bestemmingskeuzes blijkt de duur van de vakantie een belangrijke rol te spelen in de keuzes voor en tussen de verschillende buitenlandse bestemmingen. Het bezit van een toeristische accommodatie met een vaste standplaats, daarentegen, is belangrijk in de keuze voor en tussen binnenlandse bestemmingen. Verder blijkt de aanwezigheid van kinderen in het huishouden van grote invloed op de keuze voor bestemmingen dichter bij huis en is de regio of de provincie waar toeristen wonen van groot belang voor de binnenlandse regio waar men op vakantie gaat. Voor buitenlandse bestemmingen blijkt het bezit van een toeristische accommodatie zonder vaste standplaats belangrijk in de keuze vóór een buurland, speelt de planningsprioriteit van de vakantie een belangrijke rol in de keuze tussen de buurlanden, en zijn het bezit van ski’s en de keuze van het seizoen van belang voor de keuze tussen de verder weg gelegen landen. De rol van de variabelen die aangeven hoeveel dagen en/of vakanties reeds voor een bepaalde bestemming gepland zijn, blijkt verschillend voor binnen- en buitenlandse bestemmingen. Voor buitenlandse bestemmingen geven de sommatievariabelen meestal aan dat als er al een aantal dagen of vakanties in het buitenland gepland zijn, dat de kans om dat weer te doen afneemt. Voor binnenlandse bestemmingen, daarentegen, stijgt juist vaak de kans om een bepaalde bestemming opnieuw te bezoeken. Dit duidt op patronen van herhalingsbezoek binnen een bepaald jaar. Tot slot blijken de condities die “vooruitkijken” niet of nauwelijks in de beslisregels voor te komen.

Hoofdstuk 11 beschrijft de laatste drie keuzefacetten. De keuzeset voor logies-accommodaties bestaat uit (1) hotels, motels, pensions, appartementen en kamers zonder pension; (2) toeristische accommodaties in eigen bezit met een vaste standplaats; (3) toeristische accommodaties in eigen bezit zonder vaste standplaats; en (4) overige accommodaties. In de beslisregels zijn het bezit van, en de vakanties reeds gepland voor de eigen toeristische accommodaties (vast of niet) dominant. Andere belangrijke condities zijn de bestemming, de aanwezigheid van kinderen in het huishouden, de duur van de vakantie en het reisgezelschap.

De keuzeset voor het vervoermiddel bestaat uit (1) auto; (2) vliegtuig; (3) bus; (4) trein; en (5) overig. Dit keuzeproces wordt beschreven door twee beslistabellen waarbij eerst gekozen wordt tussen de auto, het vliegtuig of een alternatief vervoermiddel. Indien voor dit laatste alternatief gekozen wordt, wordt pas gekozen tussen de laatste drie alternatieven. Deze structuur is gekozen omdat de auto erg dominant is. Het is hierdoor moeilijk om minder populaire middelen te identificeren. De belangrijkste conditie in beide beslistabellen blijkt de bestemmingskeuze omdat binnenlandse vakanties niet met het vliegtuig gemaakt worden en bij buitenlandse bestemmingen vaker voor de bus gekozen wordt. Andere belangrijke condities zijn het reisgezelschap, het bezit van en het aantal reeds geplande vakanties met een auto, de logiesaccommodatie en de bestemming. Bij het afleiden van de beslistabellen zijn de “geneigdheidsvariabelen” voor vervoermiddelen ook meegenomen. Deze werden echter niet geselecteerd door het algoritme.

De keuze van de uitgaven per persoon voor een vakantie is in principe een continu keuzeproces. In *MERLIN* is echter om zowel theoretische als praktische redenen gekozen voor het discretiseren van dit proces. Hierbij worden de uitgaven aan een vakantie geoperationaliseerd als een ordinale variabele met 10 categorieën die ongeveer evenveel observaties hebben. De keuzeset voor uitgaven bestaat dus uit 10 alternatieven met elk een onder- en een bovengrens. In het simulatieproces wordt hiertussen een lineaire verdeling aangenomen. De bovengrens van het bovenste keuze-alternatief is zo gekozen dat het lineaire midden tussen de onder- en de bovengrens overeenkomt met het gemiddelde van de observaties in deze categorie. In de beslisregels bepalen de kenmerken van de vakantie meer dan voor andere keuzefacetten de uiteindelijke keuze. Het belangrijkste kenmerk van de vakantie is de bestemming: binnenlandse vakanties zijn goedkoper dan buitenlandse vakanties. Voor binnenlandse vakanties zijn verder het bezit van een toeristische accommodatie (al dan niet vast), de duur, de accommodatie en het reisgezelschap belangrijk. Voor het buitenland spelen de duur, de accommodatie en het vervoermiddel een belangrijke rol. Ook de plannings-prioriteit van een vakantie blijkt van invloed omdat aan vakanties met een hogere prioriteit vaak meer wordt uitgegeven. De aanwezigheid van kinderen in het huishouden, tenslotte, heeft voor zowel binnenlandse als buitenlandse vakanties het effect dat er per persoon minder aan een vakantie wordt uitgegeven.

Ter afsluiting van de empirische hoofdstukken brengt *hoofdstuk 12* alle bouwstenen bij elkaar. Dit hoofdstuk behandelt de laatste empirische aanvullingen en de validatie van *MERLIN*. Ook worden de toepassingsmogelijkheden van het systeem uiteengezet en gedemonstreerd. In hoofdstuk 6 is reeds aangegeven dat er gewichten geïntroduceerd moeten worden om te corrigeren voor de afwijkingen tussen de aantallen toeristische uitstapjes per jaar (uitgesplitst naar duur) in de bruikbare antwoorden op de extra vragen en in de CVO gegevens over 1998. In hoofdstuk 7 zijn namelijk de tijdallocatie modellen voor volwassenen geschat op basis van de eerste bron, terwijl aangenomen wordt dat de tweede bron betrouwbaarder is. Er worden hiervoor in hoofdstuk 12 diverse schema's van gewichten afgeleid door de simulatie-uitkomsten (na 1000 runs en op basis van de tijdallocatie modellen) te vergelijken met de CVO observaties in 1998. Voor kinderen bleken er geen gewichten nodig omdat de simulatie-resultaten de observaties perfect terugvoorspelden. Voor volwassenen, daarentegen, zijn aparte schema's afgeleid voor mensen met een toeristische accommodatie met een vaste standplaats, mensen met een toeristische accommodatie zonder vaste standplaats en voor mensen zonder toeristische accommodatie. Dit laatste bleek noodzakelijk omdat tussen deze groepen grote verschillen bestonden. De grootste gewichten waren nodig voor de keuze van "alleen dagtochten" in het participatiemodel en voor korte en extra lange vakanties in de uitstapjeskeuze modellen.

De validatie van *MERLIN* bestaat uit het vaststellen van de stabiliteit van de simulatie-resultaten en de afwijking tussen voorspelde en geobserveerde keuzes voor het standaard referentie project *MERLIN*-null. Dit referentie project simuleert (N = 200) de keuzes van de 3562 leden van het CVO-panel in 1998 zonder enige verandering in de bestaande condities. Gemeten voor ca. 100 indicatoren voor toeristische keuzes bleek de stabiliteit van de simulatie-resultaten redelijk goed. Alle indicatoren bereikten uiteindelijk een stabiele waarde en de uiteindelijke waarde (± 2.5%) van het merendeel (85-90%) van de indicatoren werd na 25-50 runs bereikt. Tevens bleken de simulatie-resultaten goed reproduceerbaar. Hiervoor werden de resultaten van het standaard referentie project *MERLIN*-null met een *t*-toets statistisch vergeleken met de resultaten van een project met identieke settings. Bij herhaling van deze toets werden steeds 3-5 significante afwijkingen gevonden, maar deze waren niet systematisch.

De afwijking tussen voorspelde (*MERLIN*-null) en geobserveerde (CVO 1998) toeristische keuzes is beoordeeld op aggregaat en op individueel niveau. In beide gevallen waren de resultaten toereikend. Op aggregaat niveau waren er echter wel grote(re) afwijkingen. Het ging hierbij met name om het aantal vakanties naar/met onbekende reisgezelschappen, de eigen toeristische accommodatie met vaste standplaats, de bestemmingen "Nederland-land-noord", Oostenrijk en Zwitserland, "verder weg gelegen bestemmingen", België en Luxemburg en Duitsland, alternatieve vervoermiddelen en de uitgaven per persoon aan korte en extra lange vakanties. Op individueel niveau bleek op basis van regressie-analyses dat het aantal extra lange vakanties, de duur van vakanties en het totaal aantal dagen dat aan vakanties besteed wordt soms meer af te wijken.

In de laatste paragraaf van hoofdstuk 12 wordt gedemonstreerd hoe MERLIN gebruikt kan worden voor scenario- en beleidsevaluaties. Allereerst kan MERLIN de effecten van re-allocatiescenarios op het niveau van het aantal toeristische uitstapjes per jaar simuleren (zie hoofdstuk 6 en 7). Als demonstratie hiervan is gekeken naar de effecten van het verhogen van het aantal vrije tijd met 12 ATV-dagen voor 10% van de werkers die minimaal 30 uur per week werken. In vergelijking met het gedrag van dit segment in 1998 blijken de extra vrije dagen gedeeltelijk te worden besteed aan vakanties, te weten korte, middellange en extra lange vakanties; het aantal dagtochten blijft contant. Ook blijkt de relatieve positie van het voorseizoen en dichtbij gelegen landen (inclusief Nederland) hierdoor te verbeteren. Andere opvallende veranderingen zijn de toename in het aantal vakanties door alleen volwassenen, reisgezelschappen met niet-schoolgaande kinderen en onbekende reisgezelschappen, het aantal vakanties in hotels en in "overige" accommodaties, en van autovakanties. Veranderingen in de bestedingen, tenslotte, komen alleen voort uit de (lichte) toename in het aantal vakanties (de gemiddelde bestedingen per vakantie per persoon blijven namelijk gelijk).

Het tweede type scenario- en beleidsevaluaties betreft de mogelijkheid om de aggregate voorkeuren voor vervoermiddelen te wijzigen (zie hoofdstuk 10). Zoals al aangegeven in de hoofdstukken 10 en 11 zijn de "geneigdheidsvariabelen" voor de vervoersmogelijkheden niet of nauwelijks geselecteerd als condities in de beslisregels voor bestemmings- en vervoermiddelkeuzes. Op basis van diverse simulaties wordt geconcludeerd dat dit type evaluaties niet gebruikt dient te worden.

Tot slot kan MERLIN scenario- en beleidsevaluaties uitvoeren door de kenmerken van de populatie te veranderen. Allereerst kan een populatie veranderd worden door de kenmerken van individuele simulatie-eenheden aan te passen. Deze methode wordt in MERLIN toegepast voor veranderingen in het bezit van toeristische accommodaties (al dan niet vast), auto's en ski's, en voor het gebonden zijn aan bepaalde (vakantie)periodes. Voor kenmerken van de bevolking die sterk samenhangen met één of meerdere andere kenmerken kan deze methode niet worden gebruikt. Dit geldt voor verandering in kenmerken zoals leeftijd, opleiding, inkomen, de werksituatie, het bezit van kinderen en het aantal leden van het huishouden. Indien bijvoorbeeld iemand in deeltijd gaat werken, dan moet ook het inkomen aangepast worden. Voor deze kenmerken wordt daarom de te simuleren populatie samengesteld door leden uit het CVO-panel uit 1998 gewogen 0, 1, 2, of meer keer op te nemen. Als demonstratie van dit principe is de vergrijzing van de bevolking tot 2020 gesimuleerd. Hierbij stijgt het aandeel van 56-plussers van 22% in 1998 tot 30% in 2020 en dalen de aandelen van de jongere groepen licht (0-15 jarigen) tot sterk (31-55 jarigen). Deze ontwikkelingen hebben, onder aanname van een constante omvang van de bevolking, tot gevolg dat de participatie aan toeristische uitstapjes sterk zal dalen. De aantallen vakanties en dagtochten dalen echter slechts licht, en de tijd besteed aan deze uitstapjes blijft zelfs constant omdat er een verschuiving optreed naar langere vakanties. Een ander opvallend effect is het toenemend belang van vakanties buiten de traditionele piekperioden en binnenlandse bestemmingen. Dit biedt bijvoorbeeld kansen voor aanbieders van het toeristisch-recreatief product om het seizoen uit te breiden en de bezettingsgraad te verbeteren. Ook is het opvallend dat de gemiddelde uitgaven per vakantie licht stijgen, maar dit komt vooral door toegenomen uitgaven in het buitenland.

Hoofdstuk 13 expliciteert de inzichten die op basis van dit onderzoek verkregen zijn. Concreet gaat het hier om vier bijdrages aan de literatuur over toerisitisch-recratieve keuzeprocessen (en daarmee aan de planning voor recreatie en toerisme) en twee methodologische innovaties. Voor de bijdrage aan de literatuur bieden de modellen die elke keuzefacet beschrijven allereerst inzicht in relaties tussen persoons- en huishoudkenmerken en de diverse facetten van het toeristisch keuzeproces. Ten tweede bieden de genoemde modellen inzicht in de relaties tussen verschillende facetten. De grootste beperking van het systeem in dit opzicht is dat – om praktische redenen – gekozen is voor een sequentieel keuzeproces, waarbij alleen keuzes die al eerder genomen zijn (in het systeem) invloed uit kunnen oefenen op latere keuzes. Het derde niveau waarop *MERLIN* inzicht biedt, is dat van de relaties tussen de toeristische uitstapjes van één persoon. Doordat in de beslisregels voor elk facet condities opgenomen zijn die weergeven welke keuzes de toerist al heeft gemaakt voor dat facet (voor andere uitstapjes), kunnen uitspraken gedaan worden over de condities waaronder mensen kiezen voor herhaling dan wel voor afwisseling. Tot slot biedt *MERLIN* inzicht in de effecten van diverse toekomst scenarios op het geheel van toeristische keuzes. Met de resultaten van dergelijke analyse is het mogelijk om de effecten van demografische en beleidsmatige trends uit te drukken in de te verwachten (relatieve) veranderingen in toeristische preferenties.

Ten aanzien van de methodologische innovaties introduceert dit proefschrift allereerst een methode om beslisregels af te leiden van empirische data. De methode is niet nieuw, maar de toepassing ervan binnen het toeristisch onderzoek is dat wel. Dit geldt ook voor de tweede methodologische bijdrage, de ontwikkeling van een simulatiesysteem waarin een groot aantal keuzefacetten, de interacties daartussen en de interacties tussen verschillende toeristische uitstapjes zijn opgenomen. Alhoewel de complexiteit van het toeristisch keuzegedrag al onderkend werd en ook technieken voor het modelleren daarvan al enige tijd bestaan zijn dergelijke complexe systemen in het toeristisch onderzoek schaars.

Kortom, MERLIN biedt interessante inzichten die beter aansluiten bij de huidige kennis van het toeristisch keuzeproces, en die daardoor een betere ondersteuning bieden voor ruimtelijke plannings- en beleidsopgaven. Alhoewel de prestaties van MERLIN redelijk zijn, vormt de huidige versie niet meer dan een eerste aanzet in de ontwikkeling van een volledig operationeel beslissings-ondersteunend systeem. Vragen voor toekomstig onderzoek hebben bijvoorbeeld betrekking op de keuze van het regel-inductie algoritme, de verdere optimalisering van de interacties tussen de verschillende keuzefacetten en uitstapjes, de wijze waarop een dergelijk systeem opgeschaald zou kunnen worden tot de hele Nederlandse bevolking en de wijze waarop beslisregels (beter) gevoelig gemaakt kunnen worden voor eigenschappen van het toeristisch-recreatief product. Desalniettemin kan op basis van dit onderzoek gesteld worden dat modellen die geen aandacht besteden aan de complexiteit van toeristische keuzes gevoelig zijn voor vertekeningen en het uitsluiten van belangrijke facetten en hierdoor misleidende beleidsconclusies tot gevolg kunnen hebben.

Curriculum Vitae

Manon van Middelkoop (1972, Voorburg) was raised in Rijswijk (South-Holland). In this 'home town' she received her primary and secondary education.

Following her secondary education (VWO) at the Lodewijk Makeblijde College (1990, cum laude) Manon replaced Rijswijk for Wageningen to become familiar with *Landscape Planning* at the Wageningen Agricultural University. Her initial interest in landscape architecture was soon to be replaced by keenness on *recreation and tourism*, environmental psychology and planning in particular. She was able to combine this focus with her interest in the English language and culture by spending some time at Loughborough University of Technology (UK).

With the degree of Master of Agriculture and Natural Environment (1996, cum laude) in her pockets, the next step was to move 'below the rivers'. In addition to acquiring a 'southern G', the objective here was to obtain a PhD. To this purpose, she was a member of the *Urban Planning Group* at the Eindhoven University of Technology from 1996 to 2001, while being employed by the Netherlands Organisation for Scientific Research (NWO).

In the near future Manon will move back to her 'home town' because she currently holds a position with the *Stichting Recreatie*, Knowledge and Innovation Centre in The Hague (since November 2000). She now uses her knowledge of academic research to advance the development of recreation in the Netherlands in the broadest sense of the word.