Specification and estimation of a logit model for housing choice in the Netherlands

Rouwendal, J.

1988

SERIE RESEARCH MEMORANDA

SPECIFICATION AND ESTIMATION OF A LOGIT MODEL

FOR HOUSING CHOICE IN THE NETHERLANDS

Jan Rouwendal

Researchmemorandum 1988-19 Juni 1988

VRIJE UNIVERSITEIT
FACULTEIT DER ECONOMISCHE WETENSCHAPPEN
EN ECONOMETRIE
AMSTERDAM

ABSTRACT

In this paper we specify and estimate a nested logit model and use it to analyze housing choice behaviour in the Netherlands. Households are assumed to decide first whether they want to continue their present housing situation or move to a rented or owner-occupied dwelling. when they choose one of the latter two possibilities, they have to choose a particular type of dwelling.

The model presented in this paper is specified in such a way that the whole decision- making process is consistent with utility maximization. Empirically it turns out that the resistance against making a move is very strong.

Since the Dutch housing market is characterized by strong and persistent excess demand and government control we have tried to incorporate the effects of the disequilibrium situation in our model. It turns out that queueing is a very important pheneomenon and that households are willing to wait for a dwelling of their most preferred type for a long period (more than a year)

In the model we have tried to deal with unobserved heterogeneity within the dwelling types that have been distinguished by using the price (rent) associated with a dwelling as an indicator of its quality.

The results of the estimation are satisfactory.

Specification and Estimation of a Logit Model for Housing Choice in the Netherlands*.

1 Introduction

In this paper we will specify a nested logit model and use it to analyze housing choice behaviour. The model is estimated for data concerning the Dutch Rimcity. Since there have been many attempts in recent years to use discrete choice models, and especially the logit model, for the analysis of housing choice (see Clark and Van Lierop[1986] for a review of this research) we will in this introduction mention a few points in which our analysis seems to differ from other approaches that have been reported in the literature. We will mention three points.

First, we have tried to build a model that can - as a whole - be considered as being consistent with utility maximizing behaviour. In this we differ from many others who have used the maximization paradigm for the analysis of the choice of the particular type of dwelling, but adopted a satisficing approach for the analysis of the decision to move or to stay. We will show that it is possible to model the whole decision process in a way that is compatible with utility maximization when the generalized costs of mobility (or : the resistance against making a move) is taken into account in a general way.

Second, we have tried to take into account the persistent state of disequilibrium of the Dutch housing market. This disequilibrium is apparent from the large number of households willing to move as compared to the relatively small number of realized moves (see VROM[1983]). The excess-demand is dealt with by means of some queueing system, which gives a priority treatment to households which are considered to be especially in need of another dwelling. The disequilibrium may be expected to have two effects. There will be some queueing for the dwelling types for which there is excess demand. On the other hand, households may decide to choose a second-best alternative when there is excess- demand for dweling type that is most preferred by them because they have to wait for a long time before they can realize a move to such a dwelling. In the model that will be specified below we have tried to incorporate these two effects.

Third, we have tried to deal with the unobserved heterogeneity that is usually present in even the finest possible classification of the housing stock and that is apparent from (large) variations in the prices of houses that do not differ in the characteristics on which the researcher has information. We will try to take this effect into account by using the price associated with a dwelling as an indication of its quality.

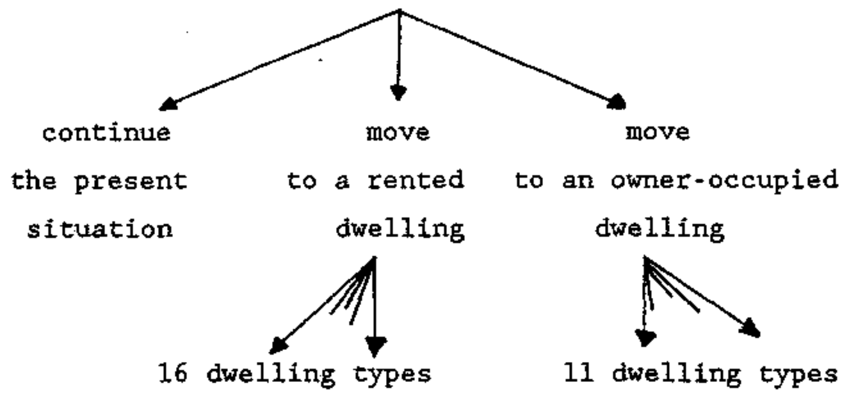
The paper is built up as follows. In section 2 we present a short overview of the complete nested model. In section 3 we deal with the unobserved heterogeneity. In section 4 we discuss the data set. In section 5 we ewxamine the adaptations to the logit model that are necessary to deal with the disequilibrium situation. In section 6 we present some figures concerning the tension between intebded and realized moves. In section 7 we specify the model and provide estimates for its coefficients. Section 8 concludes.

2 Overview of the Model.

The usual methodology for modelling choice behaviour on the housing market consists of subdividing the analysis into two stages. In the first stage the population is separated in movers and stayers, while in the second stage the destination choice of the movers is examined. In both stages of the analysis the influence of discrete choice models (logit , probit) has been profound in recent years. One of the advantages of these models is that they can be interpreted as the outcome of a utility maximizing procedure. Especially for the second stage of the relocation analysis this interpretation has been judged as being very useful. This stepwise procedure (see e.g., van Lierop[1986] and Rima and van Wissen[1988] for some recent examples concerning the Dutch housing market) is also employed in the model to be formulated in this paper.

Two stages of household decision making will be distinguished. The first one concerns the decision whether to continue living in the present dwelling or to move to a rented or an owner-occupied dwelling. In the second step the household chooses a specific dwelling type. We will try to model the decision - making in both

Figure 10.1 The Decision Tree for Housing Choices.



steps in a way that is consistent with utility maximization bahaviour. Figure 1 gives a schematic representation of the model.

The basic tool of analysis for our empirical investigation is the logit model. This model can be interpreted as the result of utility maximization when the utilities that are attached to each of the alternatives that can be chosen consist of the sum of a deterministic and a random term, the latter being (type I) extreme value distributed (see e.g., McFadden[1971]). The probability that a particular alternative n will be chosen can be written as :

$$\pi_n = \frac{e^{v_n}}{\sum_{n'=1}^{N} e^{v_{n'}}} , \qquad n=1,\dots,N , \tag{1}$$

where N is the total number of choice alternatives , and v_n is the systematic part of the utility attached to alternative n. Three submodels of the form (1) will be estimated in this paper. One referring to the choice of the dwelling type of households that are willing to move to a rented dwelling ; one referring to the choice of the dwelling type for households that are willing to move to an owner-occupied dwelling and one referring to the choice between continuation of the present situation and making a move to a rented or an owner-occupied dwelling. Some peculiarities of these submodels will be dealt with in later sections of the paper.

Throughout the chapter attention will be directed towards the understanding of the influences of allocation variables - prices and rationing - on the choice behaviour of individual households. This implies that - in comparison - less attention is devoted to other explanatory variables which are of potential importance. For instance , for the number of rooms and the type of dwelling (single family unit or apartment) only simple specifications will be used. Other variables , e.g. , the age of the head of the household (which may be expected to have some influence on the propensity to move) and the region in which one looks for a dwelling (which may restrict the set of available dwelling types : in the centre of a large city single family dwellings are hard to find) are completely absent in our analysis. It should therefore be clear from the outset that we are not looking for the ultimate model of choice behaviour on the Dutch housing market , but seek to reach the much more modest goal of understanding the influences of the allocation variables, prices and realization probabilities, somewhat better. The relevance of this should be clear from the fact that economists stress the role of these variables , while in empirical work they usually play a minor role. This can be illustrated by three recently completed studies of the Dutch housing market. In Van Lierop[1986] price-quality and income-housing costs ratios play a role in the determination of the willingness to move. Only the former ratio plays a role in the model explaining the probabilities of actual moving for those willing to move , and neither of both play a role in the choice of the dwelling type. Also in his models the regulated character of a large part of the Dutch housing market does not explicitly influence the decision-making of the individual actors (implicitly it is of course incorporated in the discrepancies between willingness to move and actual moving behaviour). In of Rima and van Wissen[1988]'s study much attention is paid to the regulated character of the market and the discrepancies between choices and realizations this implies. However, prices and incomes are completely absent in the models for household relocation estimated by them. Finally, a recent study by Scholten[1988] adopts a vacancy chain approach in which the recruitment patterns of new residents for dwellings that have become vacant are assumed to be constant over time. This implies that neither income and prices, nor rationing mechanisms play an explicit role in his model. The usefulness of this model in situations where the prices of dwellings or the composition of the housing stock or the population of households changes significantly may therefore be doubted[1].

The fact that variables which are indicated by the theory to be of great importance for the allocation process are nevertheless neglected so often in empirical work gives rise to the conjecture that it will not be easy to incorporate these variables in models which are of practical use , i.e. , in models whose coefficients have been estimated and have the expected sign and that can be used for policy evaluation. Indeed , it turned out to be the case that careful modelling was necessary in order to arrive at interpretable results. This will become clear in the following sections.

3 Unobserved Heterogeneity of Dwelling Types.

3.1 Introduction.

In empirical research one often classifies various kinds of dwellings into a relatively small number of groups which are then viewed as homogeneoous even though this assumption is not (completely) justified. Often there still exists a considerable amount of unobserved heterogeneity within the dwelling types. This is apparent from variations in the prices that have to be paid for dwellings which are of the same type. In the present section we will look at some of the consequences of this heterogeneity within groups.

3.2 The Influence on Utility

As a start , we consider a situation in which all relevant characteristics of a dwelling are known except one. The first question we intend to answer concerns the relation between the indirect utility attached to a particular alternative n , U_n^I , and the value s of the unobserved variable. This variable , which could be any unobserved characteristic of possible relevance (e.g. , the existence of shops or schools in the neighbourhood or the age of the dwelling) i, assumed to be continuous. Suppressing all other arguments of the indirect utility function U_n^I , we may write :

$$U_n^I = U_n^I(s) \quad . \tag{2}$$

We would like to know the shape of this function.

It will be assumed that households prefer a larger value of s to a lower one. This implies that the direct utility of occupying a dwelling of type n , U_n , is increasing in s , but also that households are , ceteris paribus , willing to pay a higher rent for a dwelling with a higher value of s. This implies that the price p_n will be a function of s and that variations in the latter variable can be observed indirectly by the researcher as variations in the value of the former :

$$p_n = p_n(s) \quad . \tag{3}$$

The relation (3) causes a second indirect effect on utility : the rise in the price p_n associated with the rise in the value of s causes an increase in p_n that counteracts the positive direct effect of the increase in s.

In order to see what the total effect of an increase in s on the value of the indirect utility U_n^I will be , we consider a household that has already decided to choose a dwelling of type n , and is able to rent a dwelling with a value $\bar{s}$ of the unobserved variable associated with it. It will maximize its direct conditional utility function U_n :

$$U_n = U_n(\bar{s} \; , \; \underline{q}) \quad , \tag{4}$$

where $\underline{q}$ is the vector of consumption goods other than housing , subject to a budget constraint :

$$y = p_n(\bar{s}) + \underline{u}.\underline{q} \quad . \tag{5}$$

This maximization gives rise to conditional demand functions q_k :

$$q_k = q_k(y-p_n(\bar{s}) \; , \; \underline{u}) \quad , \qquad k=1,\ldots,K \quad , \tag{6}$$

which may be substituted in the conditional utility function (4) in order to arrive at a conditional indirect utility function :

$$U_n^I(\bar{s} \; , \; y-p(\bar{s}) \; , \; \underline{u}) \quad . \tag{7}$$

Assuming the values of $\underline{y}$ and $\underline{u}$ to be constant , this is the desired form (2). Let us see how the utility value U_n^I will change as a consequence of a change in the value $\bar{s}$:

$$\partial U_n^I/\partial\bar{s} = \partial U_n/\partial\bar{s} - (\partial p_n/\partial s).\sum_{k=1}^{K}\partial U/\partial q_k.\partial q_k/\partial(y-p_n) \quad . \tag{8}$$

In this equation the first term on the right-hand side embodies the direct effect of an increase in $\bar{s}$ on the utility experienced by the household, while the second one concerns the indirect effect that occurs through the decrease in the possibilities to buy other consumption goods as a consequence of the increase in the price p_n.

If the household were completely free in the choice of s, it would choose the value of this variable in such a way that the expression in (8) became equal to zero. Assuming that the marginal utilities of all goods 1,...,K and of s are positive but decreasing and that the value of $\partial p_n/\partial s$ is non-decreasing in s, we would expect that for small values of s the first term on the right-hand side exceeds the second one, while for high values of s the reverse would be the case. One would therefore expect the function $U_n^I(s)$ to be increasing for small values of s, to reach a maximum and to be decreasing afterwards. Such a function is pictured in the north-east quadrant of figure 2.

3.3 The Relation Between Utility and Rent

In the usual case (i.e. , without unobserved heterogeneity) the conditional indirect utility U_n^I is a decreasing function of the rent p_n. The upshot of the discussion above is that this relation may be disturbed as a consequence of hidden within-class heterogeneity and that the actual relation between conditional indirect utility and rent increases for relatively low values of the rent and decreases for relatively high values of the rent only. This has consequences for the specification of the utility function in empirical work. When it happens that different prices are observed for dwelling types that are the same as judged by their observed (non - price) characteristics, a formulation of the indirect utility function that is linear in the price of the dwelling may be expected to give misleading results. Then instead a non-linear (e.g., parabolic) specification should be tried.

When unobserved heterogeneity is taken into account the price plays a double role in the indirect utility function

(i) in the usual way via the budget restriction and

(ii) as an indicator of quality.

It should therefore be expected that the relation between utility and price is different for households with different incomes. Also one may wonder whether the effects of a price increase on utility can still be separated from those of a change in the unobserved quality aspect. To this question we will return in section 8. Finally, it may be remarked that we still expect to find the same parabolic form of the relation between utility and rent when there is more than one unobserved variable.

4 Discussion of the Data Set

4.1 The Housing Needs Survey

In the Netherlands a wide variety of housing market data is available , at both the local and the national levels. An important data set that has often been used for empirical research and policy is the Dutch Housing Needs Survey (WoningBehoefte Onderzoek , abbreviated as WBO). This is a large (approximately 65,000

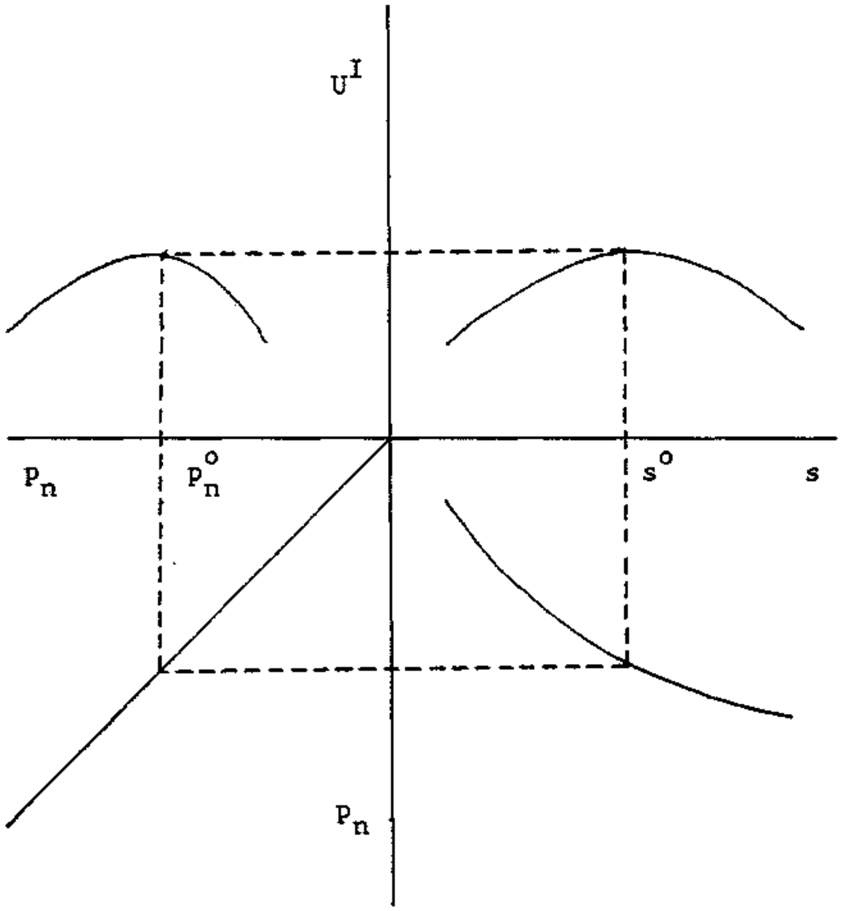


Figure 10.1 The expected relation between U^I , s and p_n.

respondents) sample of the total Dutch population that is drawn every 4 years (e.g. , in 1981 and 1985). The material under consideration here stems from 1981 (results for 1985 were not yet available). The interviewed persons were confronted with a large number of questions concerning their present and past housing situation, family conditions , age and socio-economic circumstances as well as about their plans to move to another dwelling in the near future (within 2 years). When a respondent indicated that he was considering such a move (approximately 25 % did so) , he was asked (among other questions) to what type of dwelling he would be willing to move (rented or owner-occupied), what the desired size of this dwelling would be and what price he expected he would have to pay for such a dwelling. The answers to these questions provided the basic material for the empirical work reported on below.

These answers were interpreted in the following way. A person who indicated that he was not considering a move was treated as one who viewed his present housing situation as optimal. A person who indicated he planned to move to a particular type of dwelling was assumed to have given his optimal choice alternative. Movers and stayers were assumed to have access to the same information concerning relevant circumstances on the housing market.

The price a person indicated he would be willing to pay was interpreted to be partly an indication for the quality of the dwelling concerned (see the previous section). Households with the same combination of (observed) characteristics and choosing for a dwelling of the same type (as defined by the values of observed dwelling characteristics) may nevertheless mention different prices because of (random) differences in the appreciation of unobserved characteristics.

As is clear from the discussion above , our model is estimated on the basis of the stated preferences of the households concerned and not on actually observed moves. The main reason we have chosen for this approach is the strong government intervention on the Dutch housing market in general and especially on the segment of rented dwellings. Roughly 60% of the total Dutch housing stock consists of rented dwellings and approximately 80 % of these rented dwellings are under government control, usually indirectly in the form of locally organized housing cooperations. Usually the government sets the prices every year , while local institutions ration demand where necessary. Government measures may restrict the behaviour of households so that they are forced to move to a dwelling which is , at least from their point of view , suboptimal. By concentrating on preferred moves we have tried to evade the problems that would be associated with the incorporation of these additional restrictions. This does not exclude the possibility that households take into account the situation on the housing market when stating their preferences, but we will try to deal with this phenomenon by adopting a generalisation of the logit model (see section 5).

4.2 The Classification of the Housing Stock

Housing is a heterogeneous good and it is therefore necessary to divide the housing stock into a number of classes which themselves can be considered as being approximately homogeneous. First of all, there are regional differences. For the Netherlands it is well known that these are significant (see e.g., Clark et al.[1986]), notwithstanding the small size of the country. We have tried to mitigate this problem by restricting our attention to the three western provinces North-Holland, South-Holland and Utrecht. These provinces contain the most densely populated part of the country, the so-called Rimcity. Housing market problems are mainly concentrated in this area. This is the reason why we have selected this part of the country. The administrative boundaries we have chosen are more or less arbitrary. On the one hand it may be said that they are too narrow since there is a lot of extra-regional commuting e.g., between Rotterdam and the western part of the province of North-Brabant and between Amsterdam and the province of Flevoland. On the other hand it may be remarked that the differences in the housing market conditions within the three western provinces are large, since they contain some large cities as well as areas where agriculture is the dominant activity[2].

For a classification of the housing stock within this area four dwelling characteristics were used : ownership, type (single family unit or apartment), number of rooms and price. The latter variable was used because the other three were insufficient to arrive at a classification for which the variations in the price within a class were small. Other variables (e.g., whether a rented dwelling was owned by a non-profit organisation, or whether the dwelling into which the household moves should be a new one) were not used, either because they were unavailable or because they were judged to be less relevant[3].

Ownership and type of a dwelling call for a classification into four types. The number of rooms was used for a subdivision of each of these into four more specific types of dwellings : those having one or two rooms, three rooms, four rooms or more than four rooms. For rented dwellings we distinguished furthermore dwellings with a low rent (below 250 guilders a month), those with a medium rent (between 250 and 450 guilders a month) and those with a high rent

Table 1 Classification of rented dwellings.

no.	type	number of rooms	rent	share (%)
1	single family unit	1,2,3	< 250	4.4
2	,,	1,2,3	250-450	1.6
3	,,	1,2,3	> 450	1.0
4	,,	4	< 250	8.8
5	,,	4	250-450	7.1
6	,,	4	> 450	3.3
7	,,	≥ 5	< 450	9.0
8	,,	≥ 5	> 450	5.3
9	apartment	1,2	< 250	6.8
10	,,	1,2	> 250	5.2
11	,,	3	< 250	10.2
12	,,	3	250-450	6.0
13	,,	3	> 450	2.5
14	,,	≥ 4	< 250	8.5
15	,,	≥ 4	250-450	14.0
16	,,	≥ 4	> 450	6.2

(above 450 guilders a month). For owner-occupied dwellings an analogous classification was used : dwellings with a low value (less than 125,000 guilders), a medium value (125,000-175,000) and a high value (above 175,000) guilders have been distinguished. Use of these criteria gives rise to a classification of the dwelling stock into 48 types. The number of dwellings of each type differed widely however and it was decided to cluster some of these classes together in order to arrive at a more useful classification. In tables 10.1 and 10.2 this classification is shown. In our sample 63 % of the dwellings were rented and 37 % owner-occupied.

The number of classes that have been distinguished is relatively large. It should be noted however that for classes that differ only in price we expect the coefficients of the utility functions to be the same, since we will try to deal with unobserved heterogeneity by means of the price differences (see the previous section). For this reason we have in effect distinguished 6 types of rented dwellings and 5 types of owner occupied dwellings.

Table 2 Classification of Owner-Occupied Dwellings.

no.	type	number of rooms	value	share (%)
1	single family unit	1,2,3	< 125,000	4.8
2	,,	1,2,3	> 125,000	2.7
3	,,	4	< 125,000	12.9
4	,,	4	125-175,000	10.3
5	,,	4	> 175,000	6.2
6	,,	≥ 5	< 125,000	8.5
7	,,	≥ 5	125-175,000	18.2
8	,,	≥ 5	> 175,000	23.2
9	apartment	1,2,3	all	6.9
10	,,	≥ 4	< 125,000	4.2
11	,,	≥ 4	> 125,000	2.0

5 The Consequences of Disequilibrium.

5.1 Introduction

The persistent disequilibrium on the Dutch housing market may be expected to have some consequences for the observed choice frequencies. When queueing occurs these frequencies will be biased since households that have chosen for a dwelling type for which excess-demand is particularly large have - on average - to wait longer and will therefore be overrepresented. On the other hand it may be expected that households that want to move to another dwelling within a reasonably short period avoid choosing the most heavily rationed types of dwellings and this will have opposite effects on the observed choice frequencies. In order to deal with these effects, it is likely that we have to adapt our logit model (1) to this situation. This issue will be discussed in the present section.

5.2 A Model for Discrete Choice under Uncertainty.

In Rouwendal[1988] a model is developed for the analysis of choices among a finite number of alternatives, when the realization of the alternative chosen is uncertain. This model may be relevant in markets that are not completely transparent. For instance, a household may be searching on the housing market for a particular type of dwelling without being able to find one with certainty within a limited period of time. We will use this model as the starting point for our empirical analysis of choice behaviour in the Dutch housing market.

The model is developed on the basis of three conditions. Two of these are motivated as straightforward consequences of the maximization of the expected utility of chosiing for a particular alternative, while the third is a variant of the 'independence of irrelevant alternatives' assumption which is fundamental for the logit model(see Strauss[1979]). Since we wanted to arrive at a model that would have as many of the convenient properties of the conventional multinomial logit model these conditions were considered to be appropriate.

We consider an individual decision maker (a household) that has to choose one alternative out of N possible choices. His present situation can be identified with one of these alternatives , say n^*. A choice for a particular alternative n gives a probability ψ_n that it will be realized. When the choice will not be realized the decision maker continuates his present situation. With respect to the housing market one may imagine a household that occupies a dwelling of type n^*, is willing to move to a dwelling of type n , has a probability ψ_n of realizing this desire in the current period and continues living in his present dwelling when it is unabvle to do so.

The model can be formulated as follows. The probability π_{in} that household i will choose to move to a dwelling of type n is equal to :

$$\pi_{in} = A_n \cdot \frac{e^{v_n}}{\sum_{n'=1}^{N} e^{v_{n'}}} , \qquad (9)$$

$$n=1,\ldots,N \ , \ i=1,\ldots,I \ ,$$

where I is the total number of households and A_n is a correction term which is equal to 1 when n is the index of the dwelling currently occupied by the household and equal to :

$$A_n = \frac{g(\psi_n, v_{n(i)}) \cdot \sum_{n' \neq n(i)} e^{v_{n'}}}{\sum_{n' \neq n(i)} g(\psi_{n'}, v_{n(i)}) \cdot e^{v_{n'}}} , \qquad (10)$$

$n=1,\ldots,N$, $n' \neq n(i)$,

otherwise. In this equation n(i) denotes the type of dwelling currently occupied by the household.

The model is thus basically equal to the multinomial logit model (1) , with correction terms A_n for those alternatives for which realization is uncertain. The function $g(\psi_n, v_{n(i)})$ should be increasing in ψ_n. This guarantees that alterntives with a relatively high realization probability get a higher choice probability than would be predicted by the multinomial logit model, while alternatives with a relatively low realization probability get a lower choice probability. The argument $v_{n(i)}$ gives the possibility to deal with possible effects of the initial situation on the behaviour towards risk. However, we will not use this possibility in our empirical work and shall therefore use the notation $g(\psi_n)$.

We will use the model described above for the analysis of choice bahaviour of households that are willing to move. This means that the alternative 'continuation of the present situation' can be left out of consideration. This gives us the possibility to reformulate our models as follows[4]:

$$\pi_{ni} = \frac{g(\psi_n) \cdot e^{v_n}}{\sum_{n'=1}^{N} g(\psi_{n'}) \cdot e^{v_{n'}}} , \qquad (11)$$

$n=1,\ldots,N$.

When the function $g(\psi_n)$ is specified in such a way that $\ln(g(\psi_n))$ is linear in the parameters to be estimated, the model of equation (11) can be used for empirical work in the same way as the conventional multinomial logit model.

5.3 Queueing and Uncertainty

In the present subsection we will investigate the question whether or not the model presented in 5.2 can be used as a useful approximation to the choice behaviour of individual households on the housing market. The allocation system that is functioning in the housing market is not only characterized by uncertainty about the possibility to find dwelling of the most preferred type, but also by queuing effects, related to the allocation rules used by government departments and housing corporations. We therefore have to find out whether these queueing effects can be incorporated in the model presented in 5.2.

At first sight one might be inclined to give a positive answer to the question whether that model developed can be used in situations of queueing. Uncertainty about the realization of the alternative chosen, apparent from a realization probability ψ_n smaller than one implies an average waiting time equal to $1/\psi_n$. In this way a simple translation of the model from a situation of uncertainty about the realisation of the alternative chosen to one of queueing seems to be possible.

A closer examination of this translation brings some difficulties to light. In the first place it may be expected that intertemporal considerations become of more importance in the case of pure queueing than in the case of uncertainty. In the latter case one only has to decide whether or not one should engage in a particular lottery in each period, while in the former case a decision for a number of periods (depending on the length of the queue) has to be taken. Discrepancies between choice behaviour in both situations may occur when significant changes in household characteristics have to be expected in the near future. In such a case it may be of great importance to move to another house at the right time while an immediate move would be suboptimal. Queueing allows for the possibility of timing, while stochastic rationing does not. Nevertheless, it may be expected that, in general, choice behaviour is not much different in both situations.

5.4 Effects of Queueing on the Observed Choice Probabilities

There may be another problem associated with the use of the model presented above in situations where queueing occurs. This is caused by the fact that the choice situation of an individual household is influenced by the place it occupies in a particular queue. A household becomes more inclined to stick to a particular queue when it has already joined that queue for a number of periods. A change in the alternative chosen would imply that it has to start at the end of a different queue. This would imply an overrepresentation of the number of households joining a particular queue when their behaviour is analyzed by means of the model developed in chapter 3. In that model no such effects occur since every household whose chosen alternative has not been realized will, ceteris paribus, at the start of the next period be in the same decision situation as it was at the beginning of the current period. The effects of the changes in the decision situation caused by queueing are more or less comparable to those caused by serial correlation in the random terms of the utility functions.

This effect of queueing counteracts that of avoidance of heavily rationed alternatives that is incorporated in the model developed in chapter 3. It may be reinforced by two additional possible effects. In the first place we cannot exclude the possibility that there exists serial correlation among the error terms of the utility functions. In the present context serial correlation implies that people are inclined to stick to the alternative that was once chosen, even though it was not realized. When there is no rationing households will usually be searching for one period only, after that they will remain in their new dwelling for a number of periods. In that case the problem of serial correlation can be dealt with in another way (see section 7 below). When there is rationing however, it may be expected that a number of households is searching for a long period of time and also that some of these households are persistently searching for the same alternative.

The second effect refers to unobserved heterogeneity within the groups of households (i.e., to what is sometimes called spurious state dependence , see Heckman[1980]). Consider e.g., the case in which there are two types of households and that one of these groups has a particularly strong preference for the n-th type of dwelling , which has a low realization probability associated with it while the other has only a modest preference for this type of dwelling. Assume that the researcher has been unable to differentiate between the two groups and that they are treated as one. The effect of the disequilibrium situation in the housing market will be that a relatively large number of households with a strong preference for the n-th type of dwelling will still be searching, which gives rise to a larger fraction of searching households choosing the n-th type of dwelling than would be expected on the basis of the undifferentiated model.

Furthermore, it may be the case that some actors anticipate future need by joining a queue before they are really in need of the choice alternative concerned.

It may be concluded therefore that there are strong reasons to expect effects that counteract and perhaps compensate for the avoidance of alternatives wich have low realization probabilities associated with them. This has to be kept in mind in the specification and estimation of the actual model to which we now turn our attention.

To see the consequences of these possible effect associated with queueing consider a steady-state situation in which in every period a number b of households starts searching for a particular type of dwelling and continues its search behaviour until this choice has been realized. Let us look at the consequences of this extremely consequent behaviour for the observed choice probabilities.

Let π_n be the probability that a household in the population of newly searching households prefers a dwelling of type n to all other dwellings and let ψ_n be the realization probability. It will be assumed that the population of searching households is completely homogeneous. This implies that we would expect a fraction π_n of the population to choose for alternative n in every period. (Serial correlation of the error terms is assumed to be absent).

Now consider the situation in which all these households stick to a once chosen alternative. This implies that among the new searching households the fraction preferring a dwelling of a particular type n is still π_n , but that the distribution of the other households , that have already been searching for at least one period, is influenced by the realization probabilities. In fact there is a number $(1-\psi_n).\pi_n.b$ left of households that have entered the searching population one period ago, a number $(1-\psi_n)^2.\pi_n.b$ of those that have entered two periods ago, etc. The total number of households searching for a dwelling of type n will therefore be equal to :

$$\sum_{\tau=0}^{\infty} (1-\psi_n)^{\tau}.\pi_n.b = \pi_n.b/\psi_n \quad , \qquad n=1,\ldots,N \quad . \tag{11}$$

This means that the observed fraction of households choosing a dwelling of type n , π_n^o will not be equal to π_n , but to :

$$\pi_n^o = (\pi_n.b/\psi_n)/(\sum_{n'=1}^{N} \pi_{n'}.b/\psi_{n'}) \quad , \qquad n=1,\ldots,N \quad . \tag{12}$$

This means that for choice alternatives with a relatively small realization probability the observed fraction π_n^o exceeds the 'real' fraction π_n , while for choice alternatives with a relatively large realization probability the reverse is the case.

When the assumption that all households will continue joining one and the same queue until the alternative chosen by them has been realized it becomes more difficult to adapt the model to the case of queueing. It may e.g. , be assumed that in every period a fraction α of the searchers in the queue reconsider their decision (which is taken to means that they draw a new set of random terms to their utility functions , which is independent of the old one) , while the rest simply continues joining the queue. Again assuming a fixed inflow of new households in each period we find for the observed choice probabilities :

$$\pi_n^o = (\pi_n/[1-\alpha(1-\psi_n)]/(\sum_{n'=1}^{N} \pi_{n'}/[1-\alpha(1-\psi_{n'})]) \quad , \qquad n=1,\ldots,N \tag{13}$$

which is a more cumbersome expression than (12).

It should be remarked that the effects of queueing counteract those of a'oidance of choice alternatives with a low realizatior probability as predicted by the model developed in chapter 3. This may be expected to make estimation of the avoidance effect more difficult.

5.5 Conclusion

In principle it is of course possible to model the choice behaviour of individual households in a market where queueing is relevant in a theoretically correct way, i.e., by developing the appropriate intertemporal choice model. In practice however, this route will be difficult to follow. First of all there is the problem of correlation among the error terms (these cannot be considered as independent since, during the waiting time, the household will always occupy its original dwelling. Second, one has to determine the waiting times for each of the dwelling types, which may not be an easy matter since the government may use priority schemes for various groups of searching households, while there also often exists a 'grey' market where waiting times can be much shorter. Third, this approach would make it desirable to model also the shoving process within the various queues and the differences in choice situations associated with the various places in the queue.

These arguments make clear that it would not be an easy matter to model a market where pure queueing is the allocation mechanism in the appropriate way. In this situation the best one can do seems to be to adopt the model developed in chapter 3 for the analysis of actual housing markets , even though the situation in such markets is probably not characterized by pure uncertainty about the realization of alternatives chosen but also by queueing. This is the approach that will be adopted in the rest of this paper.

6 Determination of the Realization Probabilities.

6.1 Introduction.

It has already been mentioned above that the housing market in the Netherlands , and especially the part of it that concerns rented dwellings, is characterized by persistent disequilibria. This gives reason to suspect that the realization probabilities of intended moves are smaller than one in this part of the market, i.e., that there are a number of people willing to move house, but not able to do so. This is indeed the case and one of the uses of the above mentioned Housing Needs Survey is to get an estimate of the demand for dwellings on the basis of the stated preferences of the households. In the survey households are asked whether or not they consider to move house in the near future and what would be the desired characteristics of the new dwelling.

Of course there may be doubts about the proper interpretation of the answers given to these questions. A very sceptical observer might say that a household may interpret this question as "If the land of Cocagne existed, what kind of dwelling would you prefer to live in ?". This would make the answers almost completely worthless for a trustworthy estimate of the demand for dwellings. On the other hand an observer might regard the answers to these questions as being based on proper information about the situation in the housing market , which would imply that the household is actively searching for exactly the kind of dwelling it indicates. Although the truth will probably be somewhere in the middle , we will show that the answers given in the above mentioned survey do not give the impression that the first extreme is very realistic. In the next subsections we will use the answers given to these questions in a way which comes close to the second extreme. This will not be done on the basis of a sound belief in its trustworthiness , but simply because no other way of proceeding with our analysis seemed to be possible. The fact that we have selected those households that indicated to be willing to move within a period of only one year may increase the validity of the approach.

What we would like to know is how many households that looked for a particular type of dwelling in a certain period succeeded in the realization of their intention. Since there exist no readily available figures about the number of households looking for a certain type of dwelling which are specific for the duration of search (ideally this should be longitudinal data) , we have to look for another method. What we will do is try to get an estimate of these probabilities by looking at the answers given to the questions concerning the intention to move house. We will compare these to the retrospective questions (also contained in the Housing Needs Survey) concerning moving house in the preceding three years. This gives us an indication of the number of people that moved into a dwelling of any particular type in the preceding years and the number intending to do so in the near future. The difference between these figures will presumably give us some information concerning the numerical value of the realization probabilities.

For this purpose we have used selected information on a subsample of the Housing Needs Survey held in 1981 containing a little more than 15,000 households in the three Western provinces of the Netherlands. We use the classification of the dwelling types that has been outlined in the preceding section. In the next two subsections we study the figures for the rented and the

Table 3 Realized and Intended Moves for the Rented Sector.

no.	type	rooms	rent	realized moves 1978	1979	1980	1981	intended moves
1	single family unit	1,2,3	≤ 250	21	23	18	21	42
2	,,	,,	250-450	10	7	16	10	121
3	,,	,,	> 450	6	16	16	19	44
4	,,	4	≤ 250	17	21	8	15	30
5	,,	4	250-450	36	34	28	20	189
6	,,	4	> 450	24	26	34	42	107
7	,,	≥ 5	≤ 450	19	24	23	17	95
8	,,	≥ 5	> 450	29	39	40	33	79
9	apartment	1,2	≤ 250	52	51	54	45	53
10	,,	1,2	> 250	33	39	63	69	105
11	,,	3	≤ 250	48	50	57	31	91
12	,,	3	250-450	39	51	46	47	211
13	,,	3	> 450	11	23	38	34	47
14	,,	≥ 4	< 250	35	26	35	21	45
15	,,	≥ 4	250-450	81	67	73	64	150
16	,,	≥ 4	> 450	40	33	59	47	65
				501	530	608	535	1,474

owner-occupied parts of the market.

6.2 The Rented Sector

For the 17 types of rented dwellings that have been distinguished the number of realized moves in 1978 , 1979 , 1980 and 1981 and the number of intended moves within a year have been listed in table 10.2.

The first remark that may be made about this table is that the number of moves into the various types of dwellings does not show a clear trend over time. The only exception to this general impression are the small apartments with a rent exceeding 250 guilders a month for which the number of moves increases steadily. The general impression is that neither the total number of movers into a rented dwelling, nor the distribution of this number over the various types shows a clear development over time. This seems to confirm empirical evidence given by Scholten et al.[1986] (see also Scholten[1988]) concerning the pattern of mobility on the Dutch housing market.

It is also clear from the table that the number of intended moves is much larger than the number of realized moves. The figures about intended moves give little ground for the opinion that the households answered the questions about intended moves without any consideration of the existing situations on the housing market. The largest discrepancies occur for the medium rented types of dwellings (notably for the one family units) and not for the cheapest types.

For our purposes it will be assumed that a household that indicated to be willing to move to a particular type of dwelling is

Table 4 The Realization Probabilities.

no.	prob.	no.	prob.	no.	prob.
1	0.50	6	0.30	12	0.20
2	0.10	7	0.20	13	0.55
3	0.30	8	0.45	14	0.65
4	0.50	9	0.95	15	0.50
5	0.15	10	0.55	16	0.70
		11	0.50		

indeed looking for such a dwelling and will accept such a dwelling when it receives an offer.

Furthermore it will be assumed that the number of households willing to move was approximately the same during 1978-1981 as it was at the time when the Housing Needs Survey was conducted.

On the basis of these two assumptions the realization probabilities can be approximated by taking the ratio of the number of realized moves during 1978-1981 and the number of intended moves indicated in the Housing Needs Survey. The figures presented in table 10.4 are the ratios of the average number of households moving into a dwelling of a particular type during 1978-1981 and the number of intended moves.

From table 10.4 it is obvious that the strongest rationing occurs for the medium-priced dwellings. The realization probabilities for the lowest and highest priced dwellings exceed those of the medium priced dwellings. Extremely small realization probabilities were found for the small and medium sized single family units with rents between 250 and 450 guilders a month. This may be regarded as an indication of strong disequilibria on the Dutch housing market. On the other hand however , this result may indicate some lack of reliability of the answers of the interviewees : a realization probability of only 10 % implies that on average a searching household has to look for quite a number of years before it can find a dwelling of the desired type.

From the last column of table 10.3 it is also apparent that the pattern of number of searchers for the various types of dwellings is, to some extent, the mirror-image of the pattern of the realisation probabilities : choice frequencies are in general highest for the medium-priced types of dwellings. This gives a clear illustration of what has been called 'the paradox of the housing market' (see Priemus[1984]) : the highest preferred dwellings are the most difficult to obtain. Because of this phenomenon there is a clear negative correlation between the values of the realization probabilities and the choice frequencies[5].

6.3 The Owner-Occupied Sector

The figures concerning the owner-occupied sector of the market are listed in table 10.5. They show a picture that differs a great deal from the one found for the rented sector. Between 1979-1981 the number of realized moves was steadily declining. Most probably this decline in the number of realized moves has been caused by the rapid decrease in the prices of owner-occupied dwellings around 1980 which marked the end of a long period of steadily increasing prices.

The decrease in mobility seems to occur for all types and dwellings and the distribution over the various types did not seem to change much. Scholten et al.[1986] found the same result.

Between 1978-1981 the number of realized moves exceeded the number of intended moves reported in the Housing Needs Survey. This indicates that also a decline in the number of intended moves has occurred. Only in 1981 the number of realized moves is below that of the intended moves. This may however , be regarded as a consequence of a return to a higher level of mobility after the shock of the earlier price decline and not as an indication of disequilibria.

Since we have , from an a priori point of view , reason to expect that market disequilibria are much less important for the 'free', owner-occupied sector of the market than they are for the much more regulated rented one, it was decided to set all realization probabilities for this part of the market equal to 1, implying that every household intending to move to a particular type of owner - occupied dwelling is able to do so within a period of one year.

7 Specification and Estimation Results

7.1 Introduction

As has been stated in the introduction to the present chapter a two-step procedure will be used for the analysis of housing choices of individual households. At the top of the decision hierarchy is the decision whether to remain in the dwelling presently occupied or to move to another rented or owner-occupied dwelling. The second decision concerns the choice of the exact dwelling type. The decision tree is shown in figure 1 above.

A bottom-up approach will be used for the estimation. We will

Table 5 Realized and Intended Moves for the Owner-Occupied Sector.

no.	type	rooms	value	realized moves 1978	1979	1980	1981	intended moves
1	single family unit	1,2,3	< 125	18	23	18	8	21
2	,,	1,2,3	≥ 125	13	21	14	6	10
3	,,	4	< 125	42	52	40	25	24
4	,,	4	125-175	50	43	58	29	54
5	,,	4	≥ 175	35	32	41	25	17
6	,,	≥ 5	< 125	22	20	18	8	21
7	,,	≥ 5	125-175	59	50	44	41	43
8	,,	≥ 5	≥ 175	93	79	65	32	49
9	apartment	1,2,3	all	23	55	44	29	18
10	,,	≥ 4	≤ 125	14	24	20	10	9
11	,,	≥ 4	> 125	7	8	9	9	12
				355	407	371	222	278

first estimate the choice of the type of dwelling for households willing to move to a rented dwelling and for households willing to move to an owner-occupied dwelling. Second , we will estimate the decision to buy or to rent a dwelling or to continue the present situation.

This nested approach to the estimation has the great advantage that the model can be split up in parts which are relatively easy to handle. Estimation of the complete model would concern 28 alternatives, which is large, even for logit models. On the other hand the nested procedure is known to give rise to an underestimation of the standard errors of the coefficients estimated for the higher stages of the decision process (see Amemiya[1978]). We have tried to specify the model in such a way that consistency with utility maximizing behaviour could be established. This means that we had to use the inclusive values of the utility functions estimated at the lower stage as arguments in the utility functions of the higher stage. On the one hand this has the advantage (apart from the theoretical desirability of such a structure) of stressing the relation between the various parts of the model , on the other hand the fact that this procedure does not give rise to a model for the willingness to move that can be used in its own right might be regarded as a drawback. We have chosen in favour of the theoretical arguments. Others (e.g. , Van Lierop[1986] and Rima and Van Wissen[1988]) have adopted the more practical approach.

7.2 Specification of the Utility Function for the Rented Sector

As pointed out above we start our empirical investigation by considering the people willing to move to a rented dwelling within a period of a year. The model to be used is the generalisation of the familiar multinomial logit specification of a discrete choice model dealt with in sectio 3 of the present paper.

The indirect utility function that will be used as a starting point for the empirical analysis contains four collections of arguments :

(i) household characteristics

(ii) dwelling characteristics

(iii)household income minus the costs associated with housing (i.e. , rent and costs of mobility) and

(iv) an unobserved characteristic.

Furthermore we will make use of the realization probabilities as a determinant of choice behaviour.

From the first group we use (apart from income) only family size, which will be denoted as r. From the second group we use the number of rooms (s_1) and the type of dwelling (s_2). The latter is a dummy variable , which is equal to zero for a one-family unit and equal to 1 for an apartment. Household income will be denoted - like before - by y, the rent of the n-th type of dwelling as p_n. The unobserved variable is also a dwelling characteristic and will be denoted by s_3. For the realization probabilities we use the symbol q_n. The indirect utility function associated with a move to the n-th type of dwelling of the i-th household in our sample can now be written in general form as :

$$U^I_{ni} = U^I_n(r_i \; , s_1 \; , s_2 \; , s_3 \; , y_i - p_n) \quad . \qquad (14)$$
$$n=1,\ldots,N \; , \; i=1,\ldots,I$$

The function U^I_{ni} contains only variables referring to the current period as its arguments. On the basis of the exposition given in chapter 9 it should however be clear that considerations with respect to the future are implicit in its formulation. It should be remarked also that the prices of other consumption goods and wealth have been suppressed in the present formulation. The prices of the consumption goods are (approximately) the same for all households in our sample and since we carry out a cross-section analysis the exclusion of this variable seems to be insignificant. The wealth variable should be expected to play a more important role, but since we have no data on it in our sample it was impossible to incorporate it in the analysis. However, this variable seems to be of more importance for the owner-occupied sector than for the rented sector. In the present section we will ignore it.

In order to facilitate estimation, U^I_n has to be specified as a function that is linear in its parameters. In principle a (Taylor) approximation of any desired order can be used for this purpose, but in practice one often chooses a specification that is linear in the variables or in transformations of these variables (e.g., logarithms), as well as in the parameters. We will confirm to this convention and assume that the indirect utility function U^I_{ni} of (14) is additively separable in its arguments :

$$U^I_{ni} = \alpha_{n1} + \alpha_{n2}.f_r(r_i) + \alpha_{n3}.f_{s_1}(s_1) + \alpha_{n4}.f_{s_2}(s_2) + $$
$$+ \alpha_{n5}.f_{s_3}(s_3) + \alpha_{n6}.f_y(y_i - p_n) \quad , \qquad (15)$$
$$n=1,\ldots,N \; , \; i=1,\ldots,I \quad .$$

To arrive at a further specification it may be observed that we have chosen our classification of the housing stock on the basis of the number of rooms (s_1) , the type of the dwelling (s_2) and its rent (which will, on the basis of the exposition of section 3 be used as an indicator for the unobserved variable s_3). Since these three variables appear elsewhere in the equation it seems to be better to make α_{n1} a constant which is not specific for the dwelling type, i.e., to assume $\alpha_{n1}=\alpha_1$, n=1,...,16. The coefficient α_1 cannot be estimated because of the translation invariance of the model.

Second, one may doubt whether it is the number of rooms as such that should appear in the utility function or the number of rooms related in some way to the size of the household. The latter seems to be more appropriate. We may therefore use the ratio of the number of rooms and the number of persons in the household (s_n/r_i) as an argument of the utility function , instead of both variables separately. Furthermore it seems that a low value of this variable has important negative effects on the utility attached to the particular choice alternative, but that a high value of it lacks the symmetrical large positive affects. For this reason a logarithmic transformation will be used. The coefficient that will be estimated for this variable will be assumed to have the same value for all types of dwellings.

Third, there seems to be no reason to use a transformation for the dummy variable s_2 , or to assume that its coefficient is type specific.

Fourth, the variable s_3 cannot be observed. For this reason we will make use of the hedonic price function of equation (3). It will be assumed that the price p_n is an increasing function of s_3. Inverting this relation it follows that s_3 is an increasing function of the observed price p_n. This inverted relation may be substituted into the indirect utilitiy function U^I_{ni} in order to arrive at a specification in which the price is substituted for the unobserved variable s_3. We will use two specifications, in the indirect utility function, viz., the price itself and its logarithm. It should be remarked at this point however that the substitution of an inversed hedonic price function for the unobserved variable s_3 is not without problems from an econometric point of view. To make this clear we consider the example of a linear relation between the price p_n and the unobserved quality aspect s_3 :

$$p_n = \beta'_{0n} + \beta'_{1n} \cdot s_3 \quad , \qquad n=1,\ldots,N \quad . \tag{17}$$

In this relation the term β'_{0n} is dwelling-type specific and may therefore be expected to be dependent on the kind of dwelling concerned (single family unit or apartment) and on the number of rooms. Assuming again linearity, the complete function may be specified as :

$$p_n = \beta''_0 + \beta'_{1n} \cdot s_3 + \beta'_{2n} \cdot s_1 + \beta'_{3n} \cdot s_2 \quad , \qquad n=1,\ldots,N \quad , \tag{18}$$

where all coefficients β are positive. "Inverting" this relationship by writing the unobserved variable s_3 on the left-hand-side we obtain the following equation :

$$s_3 = \beta_0 + \beta_{2n} \cdot s_1 + \beta_{3n} \cdot s_2 + \beta_{4n} \cdot p_n \quad , \qquad n=1,\ldots,N \quad , \tag{19}$$

where β_0 , β_{2n} and β_{3n} are negative and β_{4n} is positive[6]. Substitution of this relationship in the indirect utility function (20) does not only result in a positive relation between the value of indirect utility and the price of the dwelling concerned , but also in a negative relation betweeen this value and the variables s_1 and s_2. This implies that the sign of the coefficient for the variable s_2 becomes ambiguous as a consequence of the substitution of the inverted indirect utility function. The consequences for the variable s_1 may be expected to be of smaller importance because it is incorporated in transformed form only.

When the hedonic price function is specified in another way similar problems may be expected. Since there seems to be no way of avoiding them or to distinguish between the pure effect of the variables s_1 and s_2 and those introduced by the substitution of the inverted hedonic price function we will simply stick to the specification (20) after substitution of this relation. The coefficient for the variable p_n (or its logarithmic transformation) will be allowed to vary with the dwelling type (i.e. , with the variables s_1 and s_2). This may be expected to mitigate the problems to some extent.

Finally we have income minus rent as an argument of the indirect utility function. We will incorporate this value also after logarithmic transformation. This formulation implies that a change in price or income becomes less important for the utility value attached to a particular type of dwelling when the income net of housing costs is already large and also that differences in rent become a less important determinant of housing choice when income is large. The coefficient for this transformed variable will be assumed to be the same for all dwelling types.

Summarizing, we arrive at the following specification of the indirect utility function :

$$U^I_{ni} = \gamma_0 + \gamma_1 . \log(s_1/r_i) + \gamma_2 . s_2 + \\ + \gamma_{3n} . p_n + \gamma_4 . \log(y_i - p_n) \quad , \quad (20) \\ i=1,\ldots,I \ , \ n=1,\ldots,N \quad ,$$

where instead of p_n we will also use $\log(p_n)$. We expect γ_1 and γ_3 to be positive and γ_2 to be negative. The coefficient γ_0 is not identified[6]. With respect to the sign of γ_2 we should make a reservation however for two reasons. First, it may happen that the substitution of the inverted hedonic price function causes a reverse in the sign of this coefficient. Second, the western part of the Netherlands contains four relatively large cities. People who have chosen to live there restrict their choice set since in these cities apartments are of far more importance than in the Netherlands, or even its three western provinces as a whole. The result may be a larger observed choice frequency for apartments.

7.3 Effects of Uncertainty and Queueing

The rented segment of the housing market appears to be heaviliy rationed (see section 10.3). The persistent state of disequilibrium may be expected to have two counteracting effects on the observed choice frequencies. First individual decision makers will be inclined to avoid the most heavily rationed choice alternatives in favour of the ones that are less rationed. Second , queueing may be expected to occur.

In order to deal with the first effect we will use the model that has been developed in chapter 3 will for the analysis of choice behaviour of households that intend to move to a rented dwelling. Since only movers are concerned here we can use the simple form of the model given in ( ), which will be rewritten as :

$$\pi_{ni} = \frac{g_i(\psi_n).e^{U^I_{ni}}}{\sum_{n'=1}^{N} g_i(\psi_{n'}).e^{U^I_{n'i}}} \quad . \qquad (11)$$

$$i=1,\ldots,I \ , \ n=1,\ldots,N$$

This choice probability specification can be dealt with by choosing a particular specification for $g_i(.)$ and subsequently estimating $\ln(g_i(.))+U^I_{ni}$ as if this were the utility function. The estimation of the generalized logit model will then be possible in the same way as that of the usual logit model. We will make use of two possible specifications. The first specification that will be used is a simple power function of the realization probability :

$$g_i(\psi_n) = \psi_n^{\ \epsilon} \quad , \qquad (22)$$

$$i=1,\ldots,I \ , \ n=1,\ldots,N \quad .$$

The second is an exponential function of this probability :

$$g_i(\psi_n \ , \ U^I_{*i}) = e^{\epsilon(\psi_n - 1)} \quad , \qquad (23)$$

$$i=1,\ldots,I \ , \ n=1,\ldots,N \quad .$$

We thus have :

$$\ln(g_i(.))+U^I_{ni} = \epsilon.\ln(\psi_n) + U^I_{ni} \quad , \qquad (24)$$

$$i=1,\ldots,I \ , \ n=1,\ldots,N$$

and

$$\ln(g_i(.))+U^I_{ni} = \epsilon.\psi_n + \epsilon + U^I_{ni} \quad , \qquad (25)$$

$$i=1,\ldots,N \ , \ n=1,\ldots,N$$

respectively. In both cases the linear character of the function to be estimated is preserved. In both cases the expected sign of the coefficient ϵ is positive.

Second, we have to take into account the possibility that queueing takes place for the alternatives that are rationed. The effects of queueing counteract those of rationing-avoiding choice behaviour. We may expect the longest queues for those dwellings that are most heavily rationed. It has been pointed out at the end of the previous chapter that queueing may be caused by deterministic variation (instead of the usually assumed stochastic instability) as a cause for the randomness of preferences, by unobserved heterogeneity among the households and by anticipating behaviour. In order to investigate the consequences of queueing we have to make assumptions about the way in which queueing takes place. Two special cases have been treated in section 5. It was first assumed there that a choice for a particular type of dwelling would be maintained until it was realized, no matter how long the waiting time would be. On the basis of this assumption it was shown that in a stationary state the observed choice frequencies π^o_n would be equal to :

$$\pi^o_n = (\pi_n/\psi_n)/(\sum_{n'=1}^{N} \pi_{n'}/\psi_{n'}) \quad , \qquad n=1,\ldots,N \quad , \tag{26}$$

where π_n denotes the probabilty that an actor joining the queue chooses alterntive n. In this case it is easy to deal with the effect of queueing : instead of the function $\ln(g_i(.))+U^I_{ni}$ (see (24) and (25)) one estimates :

$$-\ln(\psi_n) + \ln(g_i(.)) + U^I_{ni} \quad , \qquad n=1,\ldots,N \quad . \tag{27}$$

In case of specification (22) this would simply imply that we expect ϵ to be biased by a value -1.

When a fraction α of the people in a particular queue continues searching, while the others reconsider their choices (see section 5) we have to estimate :

$$- \ln[1-\alpha(1-\psi_n)] + \ln(g_i(.)) + U^I_{ni} \quad , \qquad (28)$$
$$n=1,\ldots,N \quad ,$$

where α has to lie between 0 and 1. When $\alpha(1-\psi_n)$ is very close to zero first logarithmic term may be approximated as $-\alpha(1-\psi_n)$. In case specification (23) is chosen we expect to estimate $\epsilon-\alpha$ instead of ϵ as the coefficient before ψ.

When we adopt the still more general hypothesis that α may change with the length of the waiting time the relation between the theoretical (i.e., relevant for new searchers) and the observed choice frequencies becomes more complicated. Therefore this possibility will not be taken into consideration here.

Another approach would be to eliminate the effects of queueing by concentrating on households that have been searching for a short period only. One would expect that among these people the effect of queueing is negligible. The drawback of this approach is obviously that it restricts the number of observations. For our estimations we have used both approaches.

7.4 Results of Estimation

We estimated utility function (20) with prices both taken up untransformed and logarithmic , corrected for possible effects of uncertainty (about the realization of the alternative chosen) by means of specifications (22) and (23). Since it seemed to be impossible to estimate the effects of queueing independent from those of avoidance of heavily rationed alternatives it has been assumed that the effects of queueing would be incorporated in the empirical estimates of ϵ. For this reason ϵ would be expected to be greater than -1 both in case of specification (24) and (25), where a negative value of ϵ would indicate that the effects of queueing were more important than those of avoidance of heavily rationed alternatives. It should be remarked that ϵ could become smaller than 1 when specification (23) is adopted and $\alpha.(1-\psi_n)$ is not close to zero.

Of the four specifications the one with the prices incorporated after logarithmic transformation and with specification (22) used

Table 10.5 Estimation Results for the Rented Sector.

coefficient	variable	duration of search			
		all	≤ 1 year	≤ 6 months	≤ 3 months
γ_1	$\ln(\frac{\text{no. of rooms}}{\text{no. of persons}})$	12.50	10.74	10.00	5.39
		(11.9)	(5.9)	(4.7)	(1.8)
γ_2	type of dwelling	4.79	7.90	8.34	9.34
		(3.8)	(3.3)	(2.7)	(2.1)
γ_{31}	ln(rent 1,2,3)	2.27	3.16	3.56	4.95
		(6.8)	(5.5)	(4.7)	(4.4)
γ_{32}	ln(rent 4,5,6)	1.81	2.76	3.15	4.67
		(5.5)	(4.9)	(4.2)	(4.1)
γ_{33}	ln(rent 7,8)	1.33	2.38	2.81	4.53
		(4.0)	(4.2)	(3.7)	(4.0)
γ_{34}	ln(rent 9.10)	2.52	2.76	2.88	3.74
		(8.0)	(5.2)	(4.1)	(3.6)
γ_{35}	ln(rent 11,12,13)	1.63	2.00	2.28	3.46
		(5.4)	(3.9)	(3.4)	(3.4)
γ_{36}	ln(rent 14,15,16)	1.07	1.53	1.80	3.17
		(3.5)	(3.0)	(2.7)	(3.1)
γ_4	ln(income - rent)	8.37	8.76	12.15	19.34
		(4.7)	(2.9)	(3.0)	(3.1)
ϵ	ln(real. prob.)	-0.90	-0.88	-0.47	-0.15
		(-11.4)	(-6.0)	(-2.4)	(-0.6)
number of observations		1,108	371	185	95
loglikelihood estimated model		-2,801.5	-945.3	-474.7	-244.5
2*(change in loglikelihood)		541.1	166.7	76.4	37.9

for correction gave the most convincing results. The estimates of this variant are given in table 10.5. Estimates of the other variants can be found in the appendix to this paper.

The equations were estimated for different compositions of the sample. First we did not select for the (realized) period of search, second we choose only those who had been searching for less than one year , third and fourth those who had been searching for less than 6 months and less than 3 months.

Some remarks can be made with respect to these estimates. First , the coefficient for the ratio between the number of rooms and the number of persons in the household is large and very significant. This indicates that this ratio, which should be expected to have a value that does not differ much from one, is an important determinant of housing choice. Changes in the number of persons in a household are closely related to stages in the life-cycle. These changes in household-demographic variables seem to be of great relevance for movements on the housing market.

Second, the coefficient for the dwelling type (a dummy for apartments) has not the expected negative sign. In it has already been pointed out that this may have been caused by perverse effects introduced by the substitution of the inverted hedonic price function in the model, or by the fact that relatively many people preferring to live in a city are incorporated in the sample. Third , the coefficients for the rents all have the expected positive sign. The coefficient has a higher value for smaller dwellings. This may be interpreted as saying that in general for those dwellings unobserved quality aspects (i.e., other aspects than the type and the number of rooms) are more important than for larger ones. Fourth, the logarithm of the income net of housing costs also has the expected positive sign. In principle it would have been better to take the effects of individual housing subsidy into account in the construction of this variable[8]. This might have resulted in an even more significant coefficient. Fifth , it is easy to compute that the maximum value of the utility function in the variable rent is reached when the ratio between rent and income is equal to $\gamma_{3k}/(\gamma_{3k}+\gamma_4)$, where the index k refers to the groups of housing types that have been distinguished. This implies that the maximum will be reached when the rent takes 10 - 20 per cent of total income , which seems quite realistic.

Finally there is the coefficient for the realization probability which has always a negative sign , indicating that the effects of queueing are far more important than those of avoidance of heavily rationed types of dwellings. In fact the lack of finding a positive value for the coefficient ϵ , even in the sample consisting only of households that have been searching for less than three months , seems to indicate that there is no avoidance of heavily rationed types. The absolute value of the coefficient ϵ is higher when households that have been searching for a longer period are added to the sample , as it should be. The small differences between the values of this coefficient for the sample containing all households and that containing only those that have been searching for less than a year may be regarded as an indication that many household reconsider their choice after they have been searching for one year.

It can be inferred from the table that the coefficient for the ratio between the number of rooms and the number of people in the household becomes larger when households have been searching for a longer period of time, while the coefficients for rent and income minus rent become smaller. It appears that households that have been searching for a long period are willing to give in some of their financial desires, while their wishes with respect to the size of the new dwelling become more pronounced.

In summary it may be said that the results of the estimation are in general satisfactory.

7.5 The Owner-Occupied Sector

For the owner-occupied sector we also used a specification of the utility function that was linear in the parameters to be estimated. The dummy for apartments and the ratio between the number of rooms and the size of the household were incorporated in the equation in the same way as was done for the rented segment of the market. The price was incorporated in the equation in a somewhat different way however. For households that move to an owner occupied dwelling from a rented one the (untransformed) price of the new dwelling would be the relevant variable. But households that owned a dwelling may regard the additional amount of money they have to pay (i.e., the difference between the prices of the old and new dwelling) as the appropriate variable for decision making. For both variables initially quality-aspects may dominate, while for higher prices the budget restriction becomes the more important variable.
Empirically, the effects of unobserved heterogeneity turned out to be very pronounced for the price difference , but not for the

untransformed price. For this reason it was decided to incorporate the untransformed price in a linear way only , while the price difference was also incorporated squared.

This suggests a specification of the following form :

$$U^I_{in} = \alpha_0 + \alpha_1 . \log(s_1/r_i) + \alpha_2 . s_2 + \alpha_{3n} . p_n + d_i . [\alpha_4 . (p_n - p^*_i) + \alpha_5 . (p_n - p^*_i)^2] \quad , \quad i=1,\dots,I \; , \; n=1,\dots,N \quad , \tag{29}$$

where d_i is a dummy variable that is equal to one when the household is owning a dwelling and equal to zero otherwise ; p^*_i is the price of the present dwelling.

An important variable that was still lacking in this specification is the wealth of the household. It should be expected that wealthier households are less senstive to the price than others. Since we had no information about the wealth of the households in our sample , it was decided to use current income as a proxy for this variable. It may be expected that for wealthier households the coefficient for the price would be smaller and also that the function in square brackets in (29) would reach its maximum at a higher price difference. For this reason it was decided to make the coefficients α_3 and α_5 dependent on income by specifying $\alpha_3 = \beta_3/y$ and $\alpha_5 = \beta_5/y$.

Another variable of potential relevance would be the interest rate for mortgage loans. The variation in this rate among the households in our cross-section sample may be expected to be small however. Since no figures about it were available in our data base it was omitted from the analysis.

The final specification for our equation was therefore determined as :

$$U^I_{ni} = \beta_0 + \beta_1 . \log(s_1/r_i) + \beta_2 . s_2 + \beta_{3n} . p_n/y_i + d_i . [\beta_4 . (p_n - p^*_i) + \beta_5 . (p_n - p^*_i)^2/y_i] \quad . \tag{30}$$

We expect β_1 and β_4 to be positive and β_2 , β_{3n} , n=1,...,5 , and β_5 to be negative. Results of estimation are shown in table 10.6. They are satisfactory. The ratio between the number of rooms and the size of the household appears again as an important argument of the utility attached to a dwelling. The coefficient for s_2 is positive , but insignificant. The coefficients for the ratio between the price of the house and household income all have the expected sign and are significantly different from zero. The same is true for the price difference variables.

Table 10.6 Results of Estimation for the Owner-Occupied Sector.

coefficient	variable	estimate	(t-value)
β_1	$\log(\frac{\text{no. of rooms}}{\text{no. of persons}})$	5.40	(4.6)
β_2	type	0.11	(0.2)
β_{31}	$price_1/income$	-9.52	(-7.1)
β_{32}	$price_2/income$	-11.81	(-10.3)
β_{33}	$price_3/income$	-13.27	(-10.3)
β_{34}	$price_4/income$	-8.79	(-5.7)
β_{35}	$price_5/income$	-14.56	(-7.2)
β_4	price difference	4.39	(13.8)
β_5	$\frac{(\text{price difference})^2}{\text{income}}$	-5.03	(-2.2)
number of observations		282	
loglikelihood		-278.3	
2*(change in loglikelihood)		795.9	

7.6 The First Stage of the Decision Process

The estimation results of the submodels for choice among the various rented and owner occupied types of dwellings are used in the estimation of the submodel for the first stage of the decision hierarchy. In this stage the households are assumed to choose one of three alternatives :

(i) continue the present situation

(ii) move to a rented dwelling

(iii) move to an owner-occupied dwelling.

The utility of alternative 1 is in principle equal to that of the dwelling type presently occupied by the household as estimated in one of the submodels discussed in the previous subsections. But since the resistance toward making a move was expected to be large (partly as a consequence of the costs associated with moving) it was decided to add a constant to this utility value in order to deal with this inertia. Furthermore it was hypothesized that the resistance against making a move would be larger for households owning the dwelling presently occupied then for those renting their present dwelling. To take this effect into account a dummy variable was introduced. The utility of continuation of the present situation was therefore specified as :

$$U_{i1} = \alpha_o + \alpha_1 . U^*_{si} + \alpha_2 + \alpha_3 . d \quad , \qquad (30)$$

where U^*_i is the utility of the dwelling presently occupied as estimated in the submodels discussed above and d is a dummy for owner-occupiers. The incorporation of the terms α_2 and $\alpha_3 . d$ introduces the kind of state dependence that was discussed in 2.4.5 in the model.

The utility of moving to a rented dwelling is equal to the highest utility value that could be reached when such a move would be undertaken. This maximum obtainable utility is given by the inclusive value of these utilities, which will be denoted as U^*_{ri} and is defined as follows :

$$U^*_{ri} = \log(\sum_{n=1}^{16} e^{U^I_{ni}}) \quad . \qquad (31)$$

We thus have for the utility of moving to a rented dwelling :

$$U_{i2} = \alpha_0 + \alpha_1 . U^*_{ri} \quad . \qquad (32)$$

The utility of moving to an owner-occupied type of dwelling is basically equal to the inclusive value for the choice alternative of this segment of the market :

$$U^*_{oi} = \log(\sum_{n=1}^{11} e^{U^I_{ni}}) \quad . \qquad (33)$$

Two additional terms were introduced for this choice alternative however. First it was hypothesized that households already owning a dwelling would be more inclined to choose again for such a dwelling if they made a move than others. For this reason a dummy variable was introduced. Second, it seems probable that owning a dwelling is most attractive for households with a high income. This motivates the introduction of income as a determinant of the utility of moving to an owner-occupied type of dwelling. We thus arrive at the

Table 7 Results of Estimation for the First Stage of the Decision Process.

coefficient	variable	estimate	(t-value)
α_1	inclusive value	0.079	(8.5)
α_2	constant (for alt. 1)	2.241	(31.7)
α_3	dummy for owning (for alt. 1)	5.202	(11.4)
α_4	dummy for owning (for alt. 3)	1.949	(5.9)
α_5	income (for alt. 3)	0.039	(3.8)

number of observations	3,461
loglikelihood	-1,319.0
2*(change in loglikelihood)	4,966.6

following specification :

$$U_{3i} = \alpha_0 + \alpha_1 . U^*_{oi} + \alpha_5 . d + \alpha_6 . y_i \quad . \qquad (33)$$

Results of estimation are shown in table 7.
The coefficient for the inclusive values and the utility of the dwelling presently owned as determined in the relevant submodel of the second stage of the decision process is positive and smaller than 1. This implies that the model is consistent with utility maximizing behaviour. All coefficients have the expected sign. It appears from the values of the estimates that the resistance against making a move is strong in view of the large values of the coefficients α_2 and α_3 as compared to that of α_1. Large differences between the inclusive values of renting or buying a(nother) dwelling are needed in order to make it attractive to move. The resistance against moving is especially strong for households that own a dwelling.

The coefficients α_2 and α_3 can be interpreted as reflecting the costs associated with making a move. These include the monetary costs of moving, but also the psychic and social costs and the efforts and time needed to make a move. It appears from table 10.7 that these costs are so high that most households would prefer to stay where they are , even if a significant increase in utility could be reached by making a move. Such potential increases in utility hardly increase the probability of moving , although the effect is statistically significant.

The coefficients α_4 and α_5 indicate that households that own a dwelling and households that earn a high income are more inclined to choose an owner-occupied dwelling than other households. This confirms our expectations.

9 Conclusion

The estimation results that have been reported on in the preceding section make it clear that the prices and the quantity rationing both have a significant influence on the size and the composition of the population of households searching for another dwelling. The effects of the prices concern the choice behaviour of these households. The effects of the quantity rationing do not seem to influence choice behaviour, but cause queueing and by means of this an increase in the number of searching households. It appears that households are willing to wait for a relatively long time (at least a year) for the realization of their preferred type of dwelling.

The estimation results that have been obtained are quite satisfactory. They show that it is perfectly possible to use economic theory as a device for modelling choice processes on the housing market.

The model for housing choice that has been specified and estimated in this chapter can be used within the broader framework of a meso-economic housing market model. In such a model the effects of changes in the prices and in the realization probabilities on the distribution of the population of households over the housing stock can be studied. Such simulation exercises would be especially useful however if we also had an empirical model for household formation and dissolution at our disposal. Starting households are a very important group of demanders on the housing market. Simulation exercises that concern only households that are already participating in the market seem to be of limited value. Since we have not developed an empirical model of household demography in the present study such a broad approach was out of the question here.

Notes

* I am grateful to Kees Gorter for assistance during an early stage of the research reported on in this paper.

1 Scholten[1988] shows however that the recruitment patterns did not change significantly during 1977-1981. This gives some justification to the use of his approach for that period.

2 Other studies have used different regional boundaries. E.g. Rima and Van Wissen[1988] used the greater Amsterdam area (including a part of the province of Flevoland), Scholten[1988] distinguished the three largest towns of the Rimcity as a separate entity.

3 It may be remarked that the boundaries chosen for the prices of owner occupied dwellings are close to thos used for some subsidies for those dwellings.

4 One may object that in (11) alternative n(i) should be left out of consideration. However, by using specification (11) we allow for the possibility that a household searches for the same type of dwelling, possibly in another part of te study area. This seems to be realistic and for this reason alternative n(i) is also included.

5 Estimation of the logit model $\pi_n=(\psi_n)^{\alpha}/\Sigma(\psi_{n'})^{\alpha}$ gives rise to a very significant estimate of α equal to -0.58 with a t-ratio equal to -14.5 , which causes a decrease in the loglikelihoo of 99.6.

6 The relation between the β's and the (β')'s is easy to determine.

7 This is caused by the translation invariance of the model.

8 The Individual Housing Subsidy is probably the most important subsidy used for the Dutch housing market. It gives an income-dependent contribution to the housing costs of the lower-income groups. See Van der Schaar[1987] for details.

Appendix Estimation Results of the Other Variants for the Rented Sector

In this appendix we present the estimation results for the other three variants of the model used in the text for explaining the choice of a dwelling type for households willing to move to a rented dwelling.

Table A1. Prices logarithmic , realization probability untransformed.

coefficient	variable	duration of search			
		all	≤ 1 year	≤ 6 months	≤ 3 months
γ_1	$\ln\left(\frac{\text{no. of rooms}}{\text{no. of persons}}\right)$	11.99 (11.5)	10.22 (5.6)	9.80 (4.6)	5.37 (1.8)
γ_2	type	3.84 (3.1)	6.13 (2.8)	7.55 (2.6)	8.94 (2.1)
γ_{31}	ln(rent 1,2,3)	2.09 (6.3)	2.91 (5.1)	3.44 (4.5)	4.92 (4.4)
γ_{32}	ln(rent 4,5,6)	1.63 (5.0)	2.53 (4.5)	3.04 (4.0)	4.64 (4.1)
γ_{33}	ln(rent 7,8)	1.17 (3.5)	2.15 (3.8)	2.70 (3.6)	4.50 (4.0)
γ_{34}	ln(rent 9,10)	2.46 (7.7)	2.76 (5.1)	2.89 (4.1)	3.77 (3.6)
γ_{35}	ln(rent 11,12,13)	1.59 (5.2)	2.02 (3.9)	2.28 (3.3)	3.49 (3.4)
γ_{36}	ln(rent 14,15,16)	1.07 (3.5)	1.59 (3.1)	1.82 (2.7)	3.21 (3.1)
γ_4	ln(income - rent)	8.46 (4.7)	9.15 (3.0)	12.30 (3.0)	19.60 (3.2
ϵ	realization prob.	-2.81 (-11.0)	-2.48 (-5.5)	-1.33 (-2.2)	-0.28 (-0.4)
number of observations		1,108	371	185	95
loglikelihood		2,807.3	949.0	475.3	244.6
2 times change in loglikelihood		529.4	159.2	75.3	37.7

Table A2. Prices Untransformed , Realization Probability Logarithmic.

coefficient	variable	all	≤ 1 year	≤ 6 months	≤ 3 months
		composition of the sample			
γ_1	$\ln(\frac{\text{no. of rooms}}{\text{no. of persons}})$	3.84 (7.3)	2.16 (2.3)	1.66 (1.3)	-0.76 (-0.5)
γ_2	type	1.96 (7.2)	2.50 (4.7)	2.27 (3.3)	2.23 (2.3)
γ_{31}	rent 1,2,3	1.10 (10.4)	1.32 (7.2)	1.34 (5.5)	1.84 (4.9)
γ_{32}	rent 4,5,6	1.07 (10.9)	1.37 (8.2)	1.36 (6.1)	1.91 (5.4)
γ_{33}	rent 7,8	0.90 (8.6)	1.34 (7.7)	1.38 (6.0)	2.08 (5.7)
γ_{34}	rent 9,10	1.50 (10.9)	1.42 (5.9)	1.22 (4.5)	1.45 (3.0)
γ_{35}	rent 11,12,13	0.91 (9.1)	1.01 (5.9)	1.03 (4.5)	1.50 (4.3)
γ_{36}	rent 14,15,16	0.81 (8.0)	1.07 (6.3)	1.07 (4.7)	1.69 (4.8)
γ_4	ln(income - rent)	17.88 (9.1)	19.28 (5.8)	20.79 (4.7)	30.83 (4.4)
ϵ	ln(real. prob.)	-1.26 (-15.4)	-1.39 (-9.0)	-1.04 (-5.1)	-0.96 (-3.3)
number of observations		1,108	371	185	95
loglikelihood		-2804.6	-931.2	-469.4	-234.7
2 times change in loglikelihood		534.8	194.9	87.1	57.5

A few comments are in order. First , when the realization probability is incorporated untransformed it is alway smaller than 1 , probably indicating that the term $a.(1-\psi_n)$ is usually too large to be a good approximation for $\ln[a.(1-\psi_n)]$. (This is confirmed by the value of ϵ close to -1 in the variant reported in the text). Second , when prices are incorporated untransformed it is easy to compute that the maximum of the utility as a function of rent is reached when the income minus rent is equal to γ_4/γ_{3k} , where the

Table A3. Prices and Realization Probabilities Untransformed.

coefficient	variable	duration of search all	≤ 1 year	≤ 6 months	≤ 3 months
γ_1	$\ln(\frac{\text{no. of rooms}}{\text{no. of persons}})$	2.84 (5.1)	0.67 (0.7)	0.59 (0.4)	-1.92 (-1.1)
γ_2	type	1.41 (5.7)	1.63 (3.5)	1.75 (2.9)	1.73 (2.0)
γ_{31}	rent 1,2,3	0.97 (9.5)	1.13 (6.4)	1.21 (5.1)	1.72 (4.8)
γ_{32}	rent 4,5,6	0.97 (10.1)	1.23 (7.7)	1.28 (5.9)	1.85 (5.4)
γ_{33}	rent 7,8	0.84 (8.2)	1.27 (7.4)	1.35 (5.9)	2.07 (5.7)
γ_{34}	rent 9,10	1.36 (9.9)	1.23 (5.1)	1.08 (3.2)	1.33 (2.8)
γ_{35}	rent 11,12,13	0.87 (8.6)	0.97 (5.6)	0.99 (4.3)	1.48 (4.2)
γ_{36}	rent 14,15,16	0.88 (8.5)	1.17 (6.7)	1.15 (4.9)	1.78 (4.9)
γ_4	ln(income - rent)	17.98 (9.2)	19.48 (5.9)	20.98 (4.5)	31.24 (4.2)
ϵ	realization prob.	-3.78 (-15.7)	-4.05 (-8.8)	-3.16 (-5.1)	-2.97 (-3.4)
number of observations		1,108	371	185	95
loglikelihood		-2,815.8	-935.8	-469.8	-234.4
2 times change in loglikelihood		512.4	185.7	86.3	58.1

index k refers to the groups of types that have been taken together. This implies that the maximum would be reached when income - rent equals 1,500 - 2,000 guilders a month , which seems to be unrealistic for the higher incomes. Third , when prices are incorporated untransformed the effect of the ratio between the number of rooms and the household size becomes less important and sometimes (for the smaller samples) insignificant. Also for this specification the coefficient for the realization probabilities becomes significantly more negative than it was when prices were incorporated after logarithmic transformation. In two cases the coefficient for the logarithm of these probabilities becomes (significantly) smaller than -1 , which is hard to explain otherwise than as the result of a specification error. The same amounts to the fact that in two cases the coefficient for the realization probability is significantly negative for the sample consisting of households that have been searching for at most three months. (Remember also the negative correlation between the realization probabilities and the choice frequencies mentioned in 5.4). These remarks indicate the reasons why we have chosen the variant in which both the prices and the realization probability are transformed logarithmically.