Paper for the European Transport Conference 2010 in Glasgow, Scotland, UK

Expect the unexpected: qualitative and quantitative tools and methods to detect unintended effects of transport policies

Authors: Justen, Andreas (DLR); Schippl, Jens (KIT); Höltl, Arne (DLR); Fleischer, Torsten (KIT)

German Aerospace Center (DLR), Institute of Transport Research Berlin
Karlsruhe Institute of Technology (KIT), Institute of Technology Assessment and Systems Analyses

Abstract

This paper is based on parts of Deliverable 2 of the European Project OPTIC (Optimal Policies for Transport in Combination, see DLR and KIT, 2010). It focuses on the development of an inventory of tools and methods for the early detection of unintended effects. Additionally, it deals with the question of when and where integrating assessment approaches is recommendable along the policymaking process. The inventory of tools and methods is done for mainly qualitative approaches (denominated as structurally open methods) and mainly quantitative approaches (denominated as structurally closed methods, more precisely EU models TRANSTOOLS and TREMOVE). A methodology is presented using examples that illustrate step by step the consideration of either structurally open or structurally closed methods focusing on the detection and the assessment of unintended effects. On the basis of these examples, recommendations are given when and where the consideration of different assessment approaches is most promising regarding the initially set policy objectives.

1 Introduction

There is a constant need for adequate and high quality assessment approaches to support policy decision-making. Manifold approaches are already available, represented by a wide range of stronger qualitatively oriented approaches to mainly quantitative analytical models. The usage of such methods and models and later the interpretation of results get extremely complex and challenging as soon as different policies are bound together in packages and inter-measure interaction appears. It is clear that the methods available can help to detect ex-ante unintended effects of policy interventions in the transport sector. But the questions remain, how should the approaches be integrated into the policymaking process in such way that early detection of adverse effects becomes possible and thus, the unexpected outcomes become expected ones? What are the strengths and limitations of tools and methods and how to overcome these? Is it possible to identify when and where during the policymaking process the integration of assessment approaches is most promising and effective?

The overall objective of OPTIC is to help identify in advance possible adverse effects of policies taken in isolation and to develop methodologies for the design and implementation of optimal policy combinations which reduce adverse effects and/or provide positive synergies. In this context a major focus is on the methods and tools available to identify and evaluate these effects, and thus to provide an inventory on relevant methods, models and methodologies available. For this purpose a taxonomy was developed which allows to differentiate between the approaches and investigate their pros and cons. The most relevant tools and methods were analysed in more detail because of the need to understand the general functionalities to later evaluate to what extent they may be helpful in addressing unintended effects. With regard to quantitative models a pre-selection was done by choosing EU transport and emission models TRANSTOOLS and TREMOVE for the evaluation. This is due to their central role as models for the assessment of policies at the EU-level. The analyses undertaken in OPTIC showed that even though approaches for mitigation are available and unknown unintended effects in policy decision-making can be reduced, uncertainty remains on the agenda and adverse effects cannot be avoided entirely.

With this paper we focus on selected issues addressed in the respective Deliverable of the OPTIC project. In Chapter 2 the focus is on the definition and categorisation of intended and unintended effects, to then introduce the taxonomy of tools developed and used in the project. Chapter 3 gives an overview on the tools and methods considered in the analysis. The following chapter 4 introduces an example of the policy packaging process which was developed in OPTIC and suggests when and where to integrate assessment approaches. As this represents a main outcome of the Deliverable, parts of the final conclusions in Chapter 5 are based on the respective example.

2 A taxonomy of effects and foresight tools

Policy interventions in the transport sector are accompanied by risks and uncertainties, which become visible in the form of unintended effects. These effects can lead to a deviation from the initial policy objective or even be counter-intentional. An ex-ante assessment of effects is always based on assumptions and simplifications. Even if risk, uncertainties and, in consequence, unintended effects cannot be fully avoided there is a broad range of methods and tools to give orientation knowledge and improve the likeliness that decisions taken today do not lead to unintended effects. For using these methods properly and for enabling a solid interpretation of their results it is crucial to take into account that different categories of unintended effects can be distinguished.

Therefore, in Deliverable 1 of the OPTIC project (TSU Oxford et al., 2010) a categorisation of unintended effects was developed which is used as reference in this paper (see Table 1). The underlying logic is that on the one hand effects known in advance can be distinguished from unknown effects. On the other hand, unintended effects do not have to be automatically adverse. As the table illustrates, they might be positive as well. The problematic categories are those where negative unintended effects occur. These effects might be adverse in relation to the original targets of a policy intervention (B1). Unintended effects could affect areas not directly related to the original target that might still be inside but could as well be outside the transport system (B2).

		Consequence dimension		
		A. Intentional	B. Non intentional	
			B1 Counter intentional	B2 Secondary
Knowledge dimension	W. 'Known'	The consequences that decision makers intended with the intervention ------- *Average fuel consumption of new vehicles is reduced; less fuel is consumed*	Counter-intentional effects that were anticipated at the time of decision ------- *Cars are driven longer and consume more fuel due to lower fuel cost/km (rebound effect); models may predict the effect*	Secondary effects that were anticipated at the time of the decision ------- *Longer distances driven lead to increase in congestion;* *models may predict the effect*
	X. 'Unknown'	Advantageous effects that are not known; serendipitous ------- *New cars inspire some people to 'green driving' lifestyles, saving additional energy*	Counter-intentional effects not known at the time of decision ------- *Car manufacturers economically challenged by the standard abandon plans to develop ultralite cars*	Secondary effects not known at the time of the decision ------- *Less public propensity to use alternative travel modes due to cheaper car travel, leading to line closures*

Table 1: Typology of unintended effects to be used for the evaluation of tools and methods.
Source: TSU Oxford, 2010

A wide range of tools and methods are relevant in terms of prospective analyses of the effects of policy interventions in the transport system. At the one hand, there are computer-based tools to quantify trends and their interrelations in order to predict the 'most likely' future. On the other hand, there are techniques based much more on dialogue and discussion, that intend to examine alternative possibilities, generate visions of desirable futures, or define 'wild

cards' (unexpected, but possible, events). In practice, most foresight exercises make use of a combination of methods, and it is an important choice which methods to apply in the context of policy assessment. However, it is not always easy to identify the most promising methods for a planning process.

Key-criteria for the categorisation of tools and methods should be their ability in detecting different types of unintended effects. Naturally, all tools and methods have their limitations in detecting effects. The limitations of different approaches is made clear by using the following abstract model: The transport system can be understood as a web with nodes and with linkages between these nodes (as a good approximation, one might think of a climbing net that can be found on children's playgrounds, see Figure 1). This web-model of the transport system illustrates well that when tackling one of the nodes, this is not an isolated phenomenon but other nodes are affected as well, via the linkages between these nodes. At the more or less blurred borderlines other systems (energy, land-use, economy) are attached and interact. Some of these nodes are known by planners and researchers, some nodes are anticipated but not exactly known, other nodes are completely unknown. When remaining in this abstract image, a policy intervention in the transport sector directly affects at least one, maybe several of these nodes. At the same time, a number of nodes are affected indirectly, via the linkages. The model illustrates that a policy intervention might lead to wide ranging effects, and some of them might only become visible after the measure has been implemented.

Figure 1: 'web of nodes': a practical example that inspired the abstract model

Foresight tools are never able to systematically reproduce the full picture, neither in scope nor in depth are they able (or even designed) to fully reproduce this web. They either cut out a certain area of the web (transport models) or, at the other extreme they provide only punctual knowledge from different areas (brainstorming, open space). EU transport and emission models TREMOVE or TRANSTOOLS are no exception in that representing a certain slice or cut-out of the web, with some selected nodes and the linkages between them. It should be noted that the original epistemic function of a model is to reduce complexity in order to get a better understanding of selected factors and linkages between these factors. Models are developed for their purposes and not for all purposes. This abstract conceptualisation illustrates that wide parts of the 'world' can not be included in transport models. So, it is not possible to detect any effects in excluded areas by applying a model. Other tools with a

different and/or broader focus are needed (could be another model or a different type of tool). On this basis, we make a general distinction between two groups of tools and methods along the following criteria: Does the structure of the method allow for a high degree in openness concerning the inclusion of parameters and linkages between parameters or is the method rather characterised by a pre-defined set of nodes and linkages between these nodes?

Accordingly, we introduce one category that is called ‘structurally open methods’ and one category called ‘structurally closed methods’. In reality, there is rather a continuum than a clear border line between these two categories. Table 2 illustrates that it is possible to define clear characteristics for both of them. This categorisation has considerable overlaps with the distinction between qualitative and quantitative approaches. However, here we prefer to use ‘structurally closed’ and ‘structurally open’ as main categories, since this openness or closeness seems to be highly important for the type of unintended effects that can be detected.

Structurally open methods	Structurally closed methods
• no fixed setting • mainly explorative • never purely quantitative, strongly shaped by qualitative elements • in principle open to detect effects beyond the system boundaries	• pre-defined setting • mainly for analyses of specific situations • more or less clear understanding of the relevant parameters and causal relationships represented by the linkages between these parameters • mainly quantitative • focus on effects inside the pre-defined system • effects outside the system can not be detected
Examples: brainstorming, open space, expert workshops, explorative scenarios	**Examples:** quantitative models, cost-benefit analyses, multi-criteria analyses

Table 2: Categorisation of tools and methods

3 Examples for structurally open and structurally closed methods

3.1 Structurally open methods

In general, structurally open methods are strongly shaped by qualitative data. Quantitative data might well play a role, but the main characteristics of these approaches is the way in which they seek to integrate knowledge of expert stakeholders and lay people in the process of policymaking. As such, they can be seen as complements to quantitative approaches, which by definition are limited in scope. The involvement of members of the public and/or stakeholders improves decisions in two aspects: a) the decisions will receive more respect from the stakeholders; b) the decisions benefit from the knowledge provided by the public and/or stakeholders. Expert knowledge is needed to inform policymaking, to give insights into fields where decision-makers lack sufficient knowledge. However, in many cases there is a huge amount of information available which could be difficult to handle. It might become challenging to differentiate between crucial or important issues and unimportant or even wrong or erroneous arguments. Structurally open approaches can help to filter and structure heterogonous information or arguments. In particular for issues that are discussed controversially amongst stakeholders, structurally open methods can help to separate facts from values and thus, help that unintended effects are not overseen. In this sense, we choose workshops and focus groups as examples and briefly introduce their characteristics:

Workshops are daily business in science and in policymaking. At first glance, it might seem somewhat trivial to consider them as a sort of foresight tool. But workshops, in general, allow the exchange of information and the generation of new knowledge by bringing expertise form different fields together in a discursive way. The can be organised in a formalised and structured way, with a strict agenda and prepared key-questions, focussed on concrete results or on testing of concepts. Workshops could also be kept rather open with a sort of brainstorming character and a loosely structured agenda. Their ability to detect unintended effects as well as their strengths and weaknesses strongly depends on the tasks to be discussed and on the background of the participants. In principle, an interdisciplinary panel should be able to roughly indicate potential unintended effects of secondary character (XB2, Table 1). In contrast, highly specialised participants from one scientific community or stakeholder group might be able to get detailed insights in WA and WB1 effects.

A **Focus Group** is a method for gaining qualitative data in form of a group discussion (see Barbour, 2007, p. 2). The group receives beforehand information in such way that participants should have a common basis for the discussion. Typically, citizens, experts or potential users of a service or a technology are selected as participants, often in form of a random selection. According to Morgan (1997, p. 2) focus groups are used as a self-contained method in studies in which they serve as the primary means of collecting qualitative data. Further, they can be used to generate survey questionnaires or to develop the content of applied programmes and interventions. Focus groups might also be used to further analyse the output of a primary method; to get a better understanding or verification of specific results of a survey or also of model calculations; or to evaluate a policy programme. Focus groups allow, for instance, learning on perceptions and attitudes of a group, to test reaction on a stimulus. Concerning the detection of unintended effects focus groups could be used in a quite explorative manner to see if any effects of a policy or a policy package are mentioned and discussed in the group (secondary effects). Results could serve as an input for further analytical tools (modelling, survey, scenarios).

3.2 Integrative approaches: Scenarios

Many foresight approaches are based on a combination of tools and methods. In general, these approaches are integrating data of different character and sources, quite often they combine qualitative and quantitative data. A typical example for integrated approaches are scenarios. Scenarios are defined by many authors as a coherent illustration of possible future situations together with pathways that might lead to these situations (Kosov and Gaßner, 2008, p. 9; Grunwald, 2002). Scenarios are not 'predicting' the future, but predictions can be part of them. The term scenario itself always implies that other futures are possible. Scenarios can be seen as an adequate method to support the process of policy packaging and for assessing the effects of the packages. They are able to take different measures or variables into account and explore or analyse the mutual interrelations between these variables. As said before, several tools might be combined in a scenario process (workshops, CBA, trend analyses, models such as TRANSTOOLS and TREMOVE, Delphi, roadmaps and others).

Scenarios can illustrate possible consequences of specific activities. They help identifying uncertainties, blind spots, contradictions or dilemmas. So, they allow the identifications of unknowns and help turning unknowns into known effects. They are an appropriate tool to get a rough understanding of where unintended effects might be hidden. The ability to detect specific unintended effects depends on the scenario methodology applied. Quantitative scenarios have the advantage of producing precise results. For instance with the application of models, a certain and limited set of variables and their interrelations is analysed. Such scenarios are able to quantify the impact of policy effects according to the variables considered. Qualitative scenarios are more open regarding number and type of variables and factors included, the system delimitations are less fixed. In particular when knowledge from different sources is combined, such scenarios offer a chance to anticipate secondary effects inside the transport system and beyond.

A typical categorisation of scenarios distinguishes between forecasts, explorative approaches and backcasting. A forecast is often based on or developed in relation to an extrapolation of trends. Forecasts might be used to illustrate known intentional and known counter-intentional effects (WA, WB; see Table 1). Normative scenarios, in contrast, take norms and values explicitly into account from the very beginning. They ask questions such as 'what is a desirable future' and 'how can we get there' (Kosov and Gaßner, 2008; Schippl and Leisner, 2009). A typical form is the backcasting approach, where targets are defined that should be fulfilled at a certain point in the future (for example an 80% reduction of CO_2 emissions in 2050). Backcasting entails designing 'images' of the future that seem to allow the achievement of targets and then describing options or pathways to get there (see Banister, 1998; Hickman and Banister, 2005). Detecting unintended effects is not an explicit aim of the backcasting approach. However, designing images of the future can help getting hints on unintended effects. Explorative scenarios elaborate on possible futures, ideally it does not matter if these futures are desirable or not. The central question is 'what happens if'. In practice, however, explorative scenarios are not fully free of normative settings. At least implicitly, the selection of key drivers and variables tends to be normative as well. Explorative scenarios appear to be the most interesting ones in the context of OPTIC; the underlying 'what-happens-if-approach' explicitly tries turning unknowns into knowns. The approach is able to support detecting and highlighting potential unintended effects of all categories (X and W, Table 1).

3.3 Structurally closed methods

In OPTIC two quantitative models were selected and analysed in more detail regarding their abilities to cope with unintended effects in transport policy making. Due to their central role in the assessment of EU policies, these were transport model TRANSTOOLS and emissions model TREMOVE. The motivation for the development of TRANSTOOLS (TOOLS for TRansport forecasting ANd Scenario testing, in the following denominated as TT) was 'to produce a European transport network model covering both freight, as well as intermodal transport, which overcomes the shortcomings of current European transport network models' (Burgess et al., 2005, p. 1). In TT transport demand is estimated applying freight and passenger transport models which allow for the estimation of origin-destination (OD) matrices by travel purposes in passenger or vehicle-types in freight transport. The estimated OD matrices can then be assigned to the respective transport networks. In this sense, TT is a classical four stage transport model (trip production, trip distribution and mode choice, trip assignment) with the objective of estimating traffic flows by mode and providing information about congestion levels and associated levels of service. The focus regarding the impact analysis of policies is on large-scale infrastructure projects and changes in the cost structures of modes, whereas TREMOVEs objective is to calculate transport related emissions. TREMOVE is based on country-specific input data about transport volumes in freight and passenger transport (TRE-part) and respective vehicle stocks by type and technological emission standards (MOVE-part). Additionally a welfare module allows the assessment of changes to the economy through modified transport costs. The model can be used – as well as TT – to assess policies with impact on transport demand. Recently, efforts were made to replace the TRE-part of TREMOVE by using modelled transport demand of TT.[1]

For the purpose of this paper and the OPTIC project it is of minor importance to describe the models and their functionalities in detail (further descriptions regarding theoretical foundations and model applications, see e.g. iTREN, 2009; van Herbruggen, 2007; Newton et al., 2009; Rich et al., 2009). With regard to their ability to detect unintended effects, it is worth to discuss briefly their strengths and limitations and the output indicators they provide. For both models applies what was already mentioned before: they represent a certain slice or cut-out of the web, here the transport system, with some selected nodes and the linkages between them. The main policy relevant output parameters delivered by TT are related to the transport system (e.g. traffic volumes in passenger-km and ton-km, modal share, congestion levels and travel times), the economy and the environment (e.g. impact on GDP and employment or fuel consumption and emissions via associated impact models). TT considers all modes during network assignment, thus the effects of congestion interaction, e.g. between trucks and cars can adequately be reproduced. Additionally the model covers the whole EU and the assessment of EU wide policies is possible. This strength at the same time causes one of the major limitations: the zoning is rather rough and differs in detail among countries. Therefore a major amount of the passenger transport demand remains within the same zone of origin and is actually not modelled.[2] Fuel consumption and emissions are estimated applying TREMOVE as well as effects of changing prices to the economic welfare (e.g. consumer and producer surplus, governmental budgeting and external costs). Behavioural changes regarding

[1] Both TT and TREMOVE were further developed in the iTREN and TENCONNECT projects (iTREN, 2009, Rich et al., 2009). The idea to apply both models conjointly and ease data exchange between them led to comprehensive improvements in their common use.

[2] To ensure that overall traffic volumes are considered to adequately reproduce congestion effects, intrazonal trips are distributed within the zone using average trip lengths and travel times, matching them according to observed traffic count measurements.

mode decisions and respectively a different transport demand by mode is estimated considering a large number of empirically observed elasticities. From a transport modelling perspective this 'top-down' approach using national statistics and manifold elasticities can be interpreted as a limitation, hence the attempt to combine TT and TREMOVE is comprehensible.

In summary both models are applied for the assessment of EU-wide transport policies and are not designed for the evaluation of policies with local or regional impact. Due to their enormous spatial coverage, the provision of adequate data bases is a challenge. For instance, to fulfil requirements of more detail in modelling transport behaviour, comprehensive survey data is required that still allow e.g. to quantify parameters of mode choice when more population types or travel purposes are considered. The same applies for the availability of reliable information on country-specific Values of Time (VOTs). This is not a specific problem of TT, but a common lack of data in transport modelling. The collection of respective data that allows for the estimation of VOTs considering different countries, travel purposes and income levels (given different transport supply qualities) seems urgent as the VOTs often work as mechanisms to 'translate' transport policies in the model application. Beside VOTs and behavioural data in general, there is also a lack of information about e.g. loading factors by vehicle types in freight transport or about capacity restrictions of passenger trains. Other rather inconspicuous data such as occupancy rates in passenger transport or the conversion factors used to translate monetary units of economic activities into goods have an enormous influence on the results.

Against this background, the decision whether or not the models are able to cope with unintended effects cannot be answered unambiguously. Policy packages can become very heterogeneous and it is likely that not all effects, intended or unintended, can be assessed using the models. To get a notion of the abilities to detect unintended effects and at the same time exemplify respective shortcomings, the following two examples may serve as an attempt.

Example 1: Charging for the use of infrastructure

The policy itself can be reproduced quite reasonably with TT, as it has an direct effect on transport costs, which are considered in the modelling of decisions (whether or not to drive, which mode to choose, etc.). Thus, first order impacts of changes in transport flows and modal shift are adequately addressed by the model. So, the first question raised about general ability may be answered positively. More problematic is the second question about unintended effects. What are unintended effects in the context of a pricing scheme? We may assume that an undesired externality of the policy is higher transport costs for users with lower budgets for mobility. This would contradict the objective of offering a high level of mobility to all people in the EU.

In this context, the policymaker would 'desire' a model that predicts who would be affected by the pricing scheme, where do people live and to what extent they are affected, e.g. denominated as the financial reduction of their budget for mobility. A model such as TT in its current version is not able to answer these very specific questions, methods such as interviews, surveys or focus groups would be needed to better anticipate the reactions of affected people. We may further assume that a secondary measure is defined that copes with the expected unintended effect and still enables to achieve the anticipated positive effects of the pricing scheme. This rather simple example (one primary policy, one unintended effect, one secondary measure to ensure acceptability) let us identify two possible ways of treating the problem: a) improving the model in such way that it may cope with the requirements, b) identify alternatives for the evaluation of the unintended effect. More generally, solution a is not feasible for all potential requirements, thus solution b is of general interest in the context of OPTIC, as alternatives to quantitative approaches come also into play, such as the methods introduced above (citizens consultation via surveys, workshops or focus groups, see 3.1).

Example 2: Subsidies for the purchase of electric vehicles
The appearance of more vehicles with electric propulsions would not be reflected directly in TT. We may assume a higher share of electric vehicles as a result of subsidies, but during the estimation of transport demand and later the assignment step no differentiation between vehicle types is made within TT. But, the policy would imply different transport costs for different vehicle types if we assume less fuel costs for electric vehicles. The impact of the policy can be assessed ex-post applying impact model TREMOVE. Here, the associated reduction in emissions (anticipated, known and intended effect) due to a significant share of electric vehicles (with less or no local emissions) would be estimated adequately. Additionally TREMOVE estimates the emissions that appear due to higher electricity demand in its fuel life cycle module.

Nevertheless, the consideration of differentiated cost structures by vehicle type in the demand model of TT – with the effect that e.g. more trips are realized by a specific vehicle type due to reduced fuel costs – would require modifications of the model structure. Additionally, known but further unintended effects most probably appear. For instance, the demand for electric energy increases and new power plants are necessary which produce additional emissions. The unintended effect would be out of the scope of TT because the effect occurs in the energy sector. Again, either additional models would be required (rather a model of the energy system would be needed to get more insights here) or alternative assessment approaches of mainly qualitative nature needed. Regarding the adoption or rejection of innovations such as electric vehicles it is well known that price structures are only one factor; in particular when it comes to cars, emotions can strongly influence the purchase decision. To learn more about such motivations, surveys or structured discussion with users would be required.

The examples introduced here already indicate that even in the case of a single policy, it is likely that different assessment approaches are required to cope with those effects that take place outside the focus of the models. Having this in mind and based on the brief overview of different methods available for the assessment of policies (structurally closed, structurally open methods), in the following chapter a fictive policy example of an EU Truck Toll is introduced to exemplify the application of methods during the policy packaging process. The objective of this exercise is to illustrate in detail when and where along the policymaking process, the integration of assessment methods seems most adequate.

4 Framework for the integration of assessment approaches: guiding principles

The main reference for the exercise is the policy packaging process developed in the OPTIC project (TSU Oxford, 2010, p. 60). Note that the policy packaging process as described here is a stylized interpretation of what is expectable to be the process in real life. Even though, the process helps to illustrate the most important aspects of policymaking and indicates interdependencies between relevant decision-making steps.

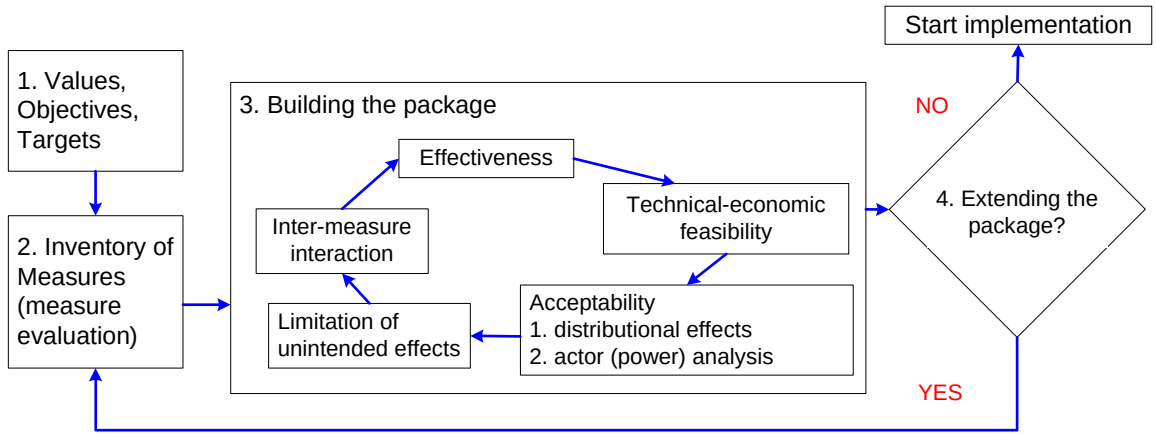

Figure 4.1: General phases in the policy packaging process – with differentiated 'Building the package part'

Starting point is the definition of an objective or target that motivates the search for and definition of policy action. In a second step a measure is defined (we assume only one primary measure is defined at the beginning). It is clear that already the steps of the definition of objectives and the selection of a measure can be very contradictory in discussion and iterative when considering different opinions from stakeholders or interest groups. The output of these steps, the definition of a measure which then needs to be assessed, is assumed as given here, thus we concentrate on the following assessment steps. This means the exercise concentrates on what is described as step 3 and 4 in the figure and the questions of: Is the policy effective, feasible, acceptable? What unintended effects may appear and how could these be assessed adequately? Are further measures necessary to achieve the policy objective and if so, how to treat inter-measure interaction? We address these issues suggesting a (sequential) methodology of how to integrate both qualitative and quantitative assessment methods in the policy packaging process to a) allow for a reliable assessment and b) to avoid the appearance of 'critical paths[3]'

At this point we skip step 1 and 2 to focus on the assessment step and directly concretize the primary measure of our example as follows: Distance and vehicle emission standard based charging scheme ('Truck Toll') for Trucks with a total weight above 12 tons and applied on national highways. For each EU member, country-specific toll values are assumed, charged by a commonly developed and EU wide introduced technology.

[3] We define ‚critical paths' as a non-desired, non-planned development of the policymaking process. A critical path leads to a deviation from the original target, objectives are not longer achieved or results only suboptimal. It is likely that a critical path leads to the appearance of (new) unintended effects.

Step 3a (see Figure 4.2) includes first aspects of the assessment of the policy. This means we enter point 3 of the policy packaging framework and the iterative process of 'Building the package'. For the detection of intended and unintended effects that come along with the proposed measure a mainly qualitative, structural open method is required (workshop, survey, focus group) to determine as many as possible of the expectable causal relationships. In the example the presented effects do not cover all effects, but rather give an indication about the variety regarding positively and negatively perceived intended and unintended effects.

The list provided distinguishes between intended and unintended effects, assuming an EU policy maker perspective. It can be expected that the qualitative evaluation done here will have a different outcome when assuming a different stakeholder perspective. Therefore, it is recommended to additionally group the intended and unintended effects by the group/stakeholder concerned with the policy impacts, thus repeat the exercise according to number and type of stakeholders included. This exercise is denominated as an actor analysis and helpful when defining additional policies with the objective to increase e.g. the acceptability of a policy (see step 4a and 4b, further below).

The list of potential effects is left open on purpose, more and other effects might be of relevance and identified according to the methodology applied for their detection. Evidently, the critical paths associated with this step are that important effects (intended, unintended) might be overseen and not considered in the further assessment steps. Therefore, it is of importance to well define the structurally open methodology applied for the detection of intended and unintended effects. The outcome of this exercise is crucial to reduce significantly adverse effects belonging to the category of 'X' (unknown intentional and non-intentional effects, see Table 1). Note that so far we are still evaluating a single primary measure.

3a. Effects	List of known intended and unintended effects

Remark: tendencies for intended and/or unintended effects depend on the stakeholder perspective – we assume an EU policy maker perspective

	INTENDED	UNINTENDED	CRITICAL PATHS
Less Truck Congestion	X		- intended and/or unintended effects are overseen and remain unknown - the methodology for the detection of intended and/or unintended effects must be chosen carefully, to assure that most likely effects are identified
Increase in Freight Rail Transport	X		
Less Truck Emissions	X		
Increase in Rail Noise		X	
Increase in Truck Flows on Secondary Roads		X	
Toll Revenues	X		
General Decrease in Demand for Trucks		X	
Higher Demand for new Trucks	X		
Increased export of old Trucks in Non-Member-States		X	
Increase in Private Car Flows		X	
Increase in Demand for local Products	X		
More Air Transport within the EU		X	
High Costs for Charging Technology		X	
Export of Charging Technology	X		
Improved Congestion Management	X		
...	...	...	

Figure 4.2: Integration of assessment approaches - step 3a[4]

[4] A table as presented here can increase transparency and enhance discussions. If respective stakeholders cannot agree on what should be in or left out of the list, several more tables can be created and used to develop different scenarios.

Although step 3a is expedient to reduce unknown effects, convert them into known effects, and make these 'visible' for the ongoing assessment, unknown intended and unintended effects remain on the agenda. The objective is to identify most of the expectable effects ex-ante, in reverse, unknown ex-post effects cannot be assessed by any approach considered here as they are unknown at the time of evaluation and decision-making. This means, although some effects of type 'X' can be converted into known effects by the exercise done before, still the existence of unknown 'background issues' should at least be recognized.

In the next step (see Figure 4.3), for the assessment of intended and unintended effects we make use of the consequence versus knowledge dimensions provided in Table 1 (WA, WB1, WB2), in such way that every effect is assigned to one of the evaluation categories. Note that only known effects are considered, as we assume we already minimized the unknown effects by the exercise conducted in step 3a. We now expand this step by the consideration of the different assessment approaches available. Note that in the example we assume TT and TREMOVE as representatives for structurally closed, mainly quantitative methods as well as the set of structurally open approaches denominated with 'SO'. This proceeding leads to a classification of intended and unintended effects regarding their ability of being assessed by the types of assessment approaches considered.

This exercise is meaningful as it gives an indication whether or not primary effects can be modelled and to what extent further structurally open methods are necessary. Additionally and with special concern to the unintended effects, the exercise allows getting a notion of what further, secondary measures are required to mitigate non-desired side effects of the policy. This step is crucial as important aspects are addressed here: (1) the definition of the type of intended and unintended effect (CTC); (2) the assignment of an assessment method (Assessment); (3) the assessment itself (which is not explained in detail) and thus, the evaluation regarding effectiveness, feasibility and acceptability; (4) finally the deduction of which unintended effects most probably need to be mitigated by secondary measures.

So far one primary measure is analyzed along the suggested framework, hence we did not yet consider secondary, additional measures. Naturally, the evaluation process will get more ambitious if we assume that various policies are introduced at the same time. Regarding the policy packaging framework this refers to the loop implied by step 4 'Extending the package' (see Figure 4.1). Looking at the list of intended and unintended effects, in the example, 7 unintended effects are expected, for those mitigation actions are required. The exercise also allows detecting effects which are somehow in-between the categories. The expected 'General Decrease in Demand for Trucks' seems an understandable result when reducing truck-based transport. In this sense, it could be denominated as an intentional effect. But at the same time production for new trucks declines, respectively do the job capacities necessary. This effect might be unintended and eventually influences the policy setting.

3b. Effects	List of known intended and unintended effects: Evaluation & Assessment

	INTENDED	UNINTENDED	CTC*	Assessment**
Less Truck Congestion	X		WA	TT
Increase in Freight Rail Transport	X		WA	TT
Less Truck Emissions	X		WA	TM
Increase in Rail Noise		X	WB2	TM
Increase in Truck Flows on Secondary Roads		X	WB1	TT
Toll Revenues	X		WA	TT
General Decrease in Demand for Trucks		X	WB2	SO
Higher Demand for new Trucks	X		WA	SO
Increased export of old Trucks to Non-Member-States		X	WB1	SO
Increase in Private Car Flows		X	WB1	TT
Increase in Demand for Regional Products	X		WA	SO
More Air Transport within the EU		X	WB2	TT
High Costs for Charging Technology		X	WB2	SO
Export of Charging Technology	X		WA	SO
Improved Congestion Management	X		WA	SO
...	...	...	...	...

CTC* = Core Typology Categories (see Table 1)
Assessment**: TT = TransTools, TM = Tremove, SO = mainly structural open methods

WA = known, intentional
WB1 = known, unintended, counter-intentional
WB2 = known, unintended, secondary

CRITICAL PATHS

- assignment to one of the assessment categories can be wrong due to missing knowledge about model functionalities or ambiguity of the measure is not given

Figure 4.3: Integration of assessment approaches - step 3b

In response to the need identified in step 3b, with step 4a (see Figure 4.4) additional, secondary measures are added. These can be categorized according to their main objective of increasing effectiveness, feasibility or acceptability. Again, a classification is suggested in accordance to what type of assessment approach should be applied. First, we assign each secondary measure to one of the categories (SMType: EP, AP, and FP). Second, we check for the type of assessment approach required to evaluate both measures (Assessment). The example indicates that the outcome of this exercise results in a more heterogeneous picture of required assessment methods.

The advantage of this exercise is that complexity becomes transparent, although interaction between effects of the two measures may occur. As we now assess a policy combination, we start dealing with the aspect of 'inter-measure interaction'. It is recommended to analyze the measures first separately, to get later an indication of the inter-measure interaction when analyzing both of them together. As long as quantitative assessment is possible (TT or TREMOVE in our case), it is recommended to first analyze the effect of the single primary measure and then based on the results, conduct the combined assessment, once again, to get a clear picture of inter-measure interaction. When adding one or more secondary measures, second order intended and unintended effects can appear. In this case the analysis for the detection of intended and unintended effects introduced in step 3a should be repeated.

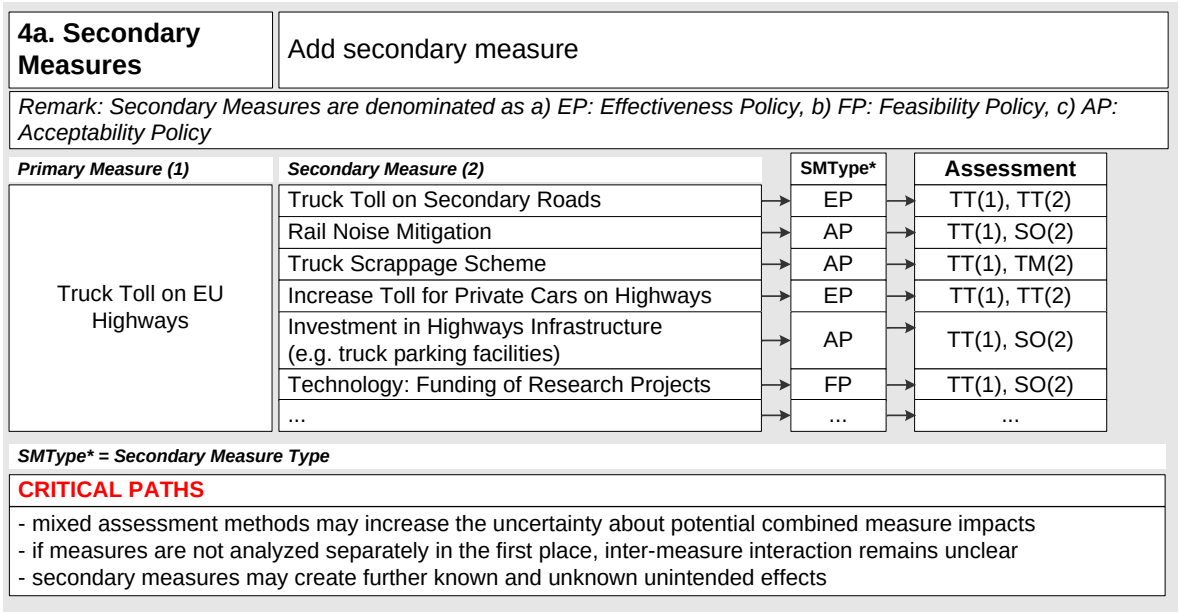

Figure 4.4: Integration of assessment approaches - step 4a

Continuing the proceeding by adding another measure leads to even more combinations of policy packages. The example of step 4b (see Figure 4.5) indicates some of these combinations. Once again we conduct the exercise of assigning the secondary measures to SMTypes and assessment methods. It is most likely that the policy combinations require for an assessment using different approaches. In the examples indicated below, the third package (Truck Toll + Investment in Highways + Technology Funding) implies two structural open assessment steps added to the quantitative assessment of the Toll using TT. Considering more measures, it gets likely that mainly qualitative assessment approaches play a significant role. Also the analysis of inter-measure interaction becomes more challenging when the effects are not evaluated by one single assessment approach. Again, additional measures create further known and unknown effects which – in the case they are perceived of having major impacts – should be identified repeating step 3a.

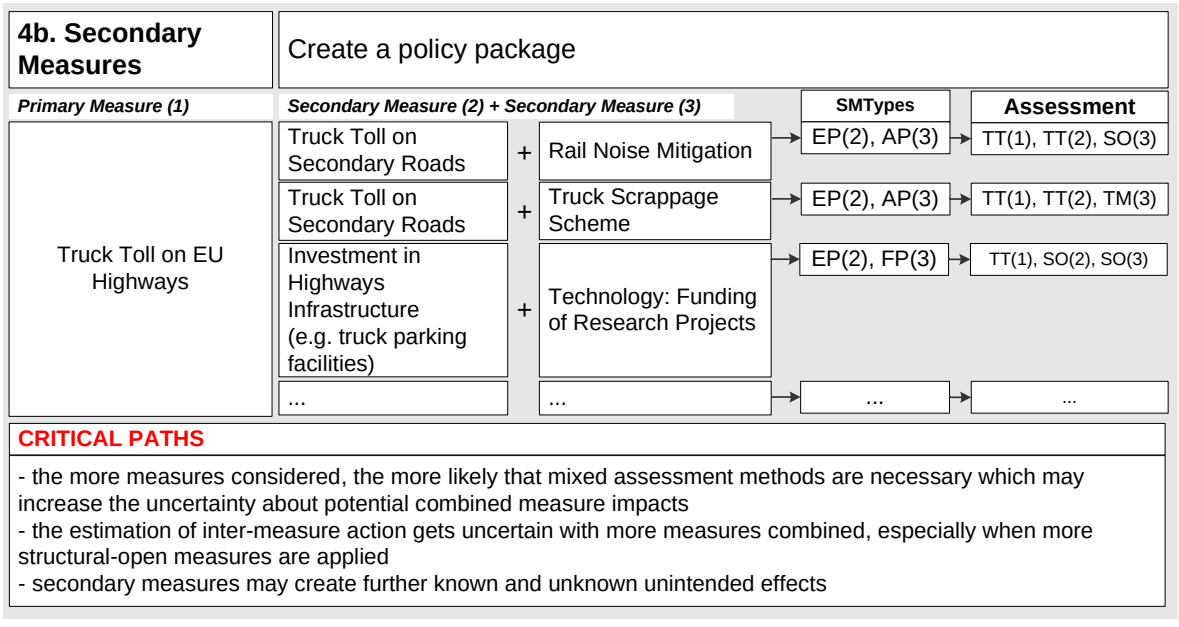

Figure 4.5: Integration of assessment approaches - step 4b

This stepwise approach introduced so far can be extended considering more policies in combination. In practice it is likewise that even more than one primary measure is applied, accompanied by several additional measures to mitigate negative, unintended effects. Following the example, there is a natural limit where the illustrations become hard to interpret because of the number of relations involved. But, the question is if all policies need to be assessed in the way described here or if there is a chance to select and/or prioritize measures. To deal with complexity, we suggest prioritizing between measures. It makes sense to apply the methodology for those measures that are expected to have a major contribution in achieving the policy goal. Small, additional measures with only a minor impact might be left out of the analysis.

In general terms, it is recommended to especially apply step 3a (identification of intended and unintended effects) and step 3b (evaluation and assessment) to additional measures of relevance. Even if a model-based ex-ante assessment is not possible in every case, step 3a will help to identify unintended effects of the policy, both negative and positive ones. The results of step 4a and 4b will support the understanding of where along the continuum from mainly quantitative to mainly qualitative assessment methods, the final package is situated. This creates awareness for the assessment methods needed. It was stated before, that it is expectable that a mix of assessment methods is required with more policies assessed in one package. As soon as 'SO' approaches play a dominant role in the overall assessment, the evaluation of the entire policy package necessarily tends to become a qualitative character.

4.1 Summary

The proceeding indicates that a policy package set up by different policy types most likely will require a multi-assessment approach. The issue of inter-measure interaction and second order intended and unintended effects (that may counteract the effects identified for the primary measure) practically impede the application of a single assessment approach. This problem remains a challenge during the assessment, but the methodology introduced before creates awareness for unintended effects and offers a structure for the identification of which assessment methods are needed. In sum, several lessons can be learned from the guiding principles introduced:

- Resources should be dedicated to identify intended and unintended effects using a qualitative approach; this exercise is decisive to reduce the unknown effects and to recognize the need for different types of assessment methods
- It is recommended to get back to the identification of intended and unintended effects after adding additional measures; possible overlaps in effects and/or counteracting impacts can be detected (at least qualitatively)
- It needs to be clearly defined which measure is analyzed by either a structurally open or structurally closed method; this should be done after having identified intended and unintended effects
- It is recommended to assess primary measures in isolation and afterwards conduct the combined assessment considering further approaches; until the policy package becomes not too complex (in the sense of illustrating all processes and interdependencies as done in the example), a notion for inter-measure interaction can be achieved
- The process helps to indicate the expected resources needed for policy evaluation; according to the number of measures that – after going through the process – are either evaluated by a model or a SO approach, required resources (model runs, stakeholder or expert consultations, etc.) can be estimated; naturally, this requires certain experience with the resources needed for the application of each assessment type

Additionally, the following remarks regarding how to operate the methodology can be given: We can assume that the definition of objectives and targets and the selection and definition of primary measures is done initially be the policymaker or policymaker group (step 1 and 2, not illustrated). It is important that the suggested methodology assumes that after step 3 the policy package is not yet fixed, even though several primary measures were defined. With step 3a a stakeholder consultation using a structurally open approach is recommended. Primary measures are discussed and the intended and unintended effects detected, considering heterogeneous perspectives on the expected effects. This step can already lead to a reformulation of the primary measure(s), if e.g. a large amount of unintended effects is detected during the stakeholder exercise. Step 3b is about the assessment of impacts, which means the application of models and/or other methods enters the proceeding. The outputs of the models are then evaluated once again either by the policymaker group (if no more stakeholder integration is desired) or in combination of policymakers and stakeholders. According to the results (effectiveness, acceptability, feasibility of the policies) and taking into account the number and type of unintended effects detected in the first place, secondary measures are defined (steps 4a, 4b). Again, each additional measure might be analysed using a structurally open approach regarding intended and unintended effects by the policymakers and/or the stakeholder group.

As a variation, it is also possible to ‘cluster’ a small number of measures in explorative scenarios. The scenarios could be built and played through with different stakeholder groups in a rather simple and qualitative process, by developing descriptions or images of how the future might look like if different measures are introduced. In addition, external factors could be varied (oil price, development in GDP, demographic changes) to assess the robustness of the planned measure in the light of different situations. It can be an explicit part of the scenario methods to separate a purely explorative or epistemic step from a quantifying evaluative step: so, in a first step the desirability of the developments are not considered; only in a second step norms and values come into play. However, such an approach is getting more resource intensive. Once such ‘starting scenarios’ are built, they can be used as a framework for the modelling phase. Results of the modelling would than be included in the scenarios and could again be discussed with stakeholder or citizens. This example illustrates that it makes sense to explicitly embed models in a broader methodological framework.

5 Conclusions

In the following we summarize and interpret the main findings. This is done along some principal conclusions that are based on the explanations and exercises realized in the previous chapters.

A clear taxonomy of tools and methods helps to structure the assessment of the approaches itself and is also useful when it comes to decide which approach to apply during the policy packaging process.

The distinction between 'structurally open' (mainly qualitative) and 'structurally closed' (mainly quantitative) allows distinguishing between approaches of very different nature. It helps to make the pros and cons of the approaches more transparent. This is relevant when they are used in the assessment of policies. Required resources, both financial and intellectual, are quite different and these differences need to be considered. Otherwise probably a non-adequate assessment is done or time running short so that no profound analysis and interpretation of results is possible.

There are no tools or methods that are able to detect ex-ante all unintended effects. Uncertainty and the possibility of unknown intended and unintended effects remain. Quantitative assessments must be understood as a necessary but in many cases not sufficient approach to reduce number and magnitude of unintended effects.

It would be negligent to assume that with the application of a model, all intended and unintended effects can be detected. If policymakers concentrate only on these aspects that are within the scope of the model or method applied, important effects (positive and negative) will be overseen. Many authors emphasize that it is highly important to make risks and uncertainties transparent and to address them in an appropriate manner. Quantitative approaches generally reproduce only a part of the 'web' of interactions, in our case of the transport system. Structurally open, qualitative approaches are, in principle, able to integrate a variety of different knowledge and thus, to anticipate previously unknown interrelations and effects. But the results are not that clear and homogenous as those produced by models.

The appropriate design and integration of structurally open methods in the policymaking process can help anticipating and thus, reducing the number of potential (unknown) unintended effects.

It is recommended by many authors and supported here that structurally open, qualitative approaches should be integrated more systematically in the process of policy making. After the definition of a policy and before starting the step of quantitative/qualitative assessment an effort on the identification of unintended effects is indispensable. The most adequate methodology to apply may depend on the specific situation (policy type, available time and resources, number and type of stakeholders), but independently of the method applied as an outcome a transparent list of expected intended and unintended effects must be available. In case of serious controversies, further discussions and/or working with scenarios might be a solution. One benefit of this is to get a more complete and transparent picture of policy impacts, another benefit is that the overview of effects serves as indication of what type of assessment approaches is now required for evaluation and/or quantification and what would be the necessary resources required. The latter point makes the question of resources explicit. A positive externality of the detection of unintended effects in advance is that resources needed for ongoing planning steps can be better estimated. A clearer picture of required resources (time, money, personal capacities) can be drawn from the number and type of unintended effects discovered.

Models can well support an ex-ante identification of unintended effects, but they need to be embedded systematically in a broader framework for analyses, including structurally open methods. Mixed approaches for the assessment of complex policy packages are recommended.

Following the policy packaging framework provided in Figure 4.1 and adding additional measures to a primary measure, complexity increases and most likely a mix of assessment approaches is required. This is not per se negative, but increases personal exigencies for a meaningful interpretation of the results. There will often be no other choice than combining methods, as models – as mentioned above – focus on a part of the 'web' of interactions. We do not vary from this position but recommend, e.g. referring to the guiding principles introduced before to analyze carefully which of the measures can be modelled, which not and what are the associated unintended effects. This most likely leads to a sort of iteration within the assessment when additional measures force us to get back to an earlier step where the qualitative detection of unintended effects is necessary. We also recommend that the analysis of unintended effects and the impact of a policy package should not start with the final package. Even though several measures are bound together to a package, a separate assessment of the primary measure – ideally using a mainly quantitative method – is recommended. If then successively further measures are added, still a notion of inter-measure interaction can be achieved. Again, it is important to at least separate the impact expected by the primary from those of additional measures. Otherwise, if inter-measure interaction is not addressed explicitly, the interpretation of the impact of a policy package becomes a somehow random exercise where every stakeholder interprets and explains the outcomes according to his/her interests.

If methods become too complex they reduce transparency of the decision-making process. A balance is needed between complexities of the methods to adequately reproduce parts of the 'web' and at the same time maintaining them understandable for non-experts in the field.

We assume that there is an interest of the decision-maker to at least conceptually understand the methods and models applied for policy assessment. The impression that a 'black box' was used for the assessment most likely will reduce the acceptance by the decision-makers as well as by those affected by the policy and thus, the likelihood that results are considered when decision are taken. The issue of transparency should also play a role when it comes to the interpretation of results. Very often underlying assumptions and used data sources are not published. Additionally, analysts that support the decision-making with their methods and tools should create the awareness that predicted developments are most likely expectable tendencies and not fixed, exact results due to the uncertainties associated with causal relationships and data.

In summary there is no unambiguous 'blueprint' for successful policy packaging considering tools and methods for the early identification of unintended effects. Nevertheless, it seems as if an ideal approach for tackling complex policy problems makes use of tools and methods by being open in the beginning of the policymaking process, closing up in the middle and being open in the end again:

- Use of structurally open methods in an explorative phase at the beginning; if several options are to be discussed, explorative scenarios could be an appropriate mean; qualitative elements are important in this starting phase, but need to be accompanied by simple pre-structured approaches (e.g. a cost-benefit-analysis) to support the exclusion of obviously unrealistic options; in this phase, stakeholders, experts as well as the wider public should be involved actively

- Using structurally closed approaches in the middle of the planning process in an analytical phase where quantifications are the main task; in this phase, work is dominated by experts and e.g. the application of analytical models

- Using structurally open, discursive methods for an interpretation of results; potentially affected groups should be included

The paper focused on tools and methods, especially their role in ex-ante assessment. We discussed options of when and how to integrate different assessment approaches aiming at an early detection of unintended effects. Naturally, the policy process is not static with a defined policy that effectively heads towards the objective during its entire time of implementation. Framework conditions evolve and most likely, the effectiveness of the measure alters over time in such way that corrective action, adjustments up to a complete policy reformulation is necessary. This is where in addition to the ex-ante assessment options discussed in this paper, further possibilities for an ex-post identification of unintended effects and potentials for remedial action need to be considered. In practice this means the extension of the principles introduced here by aspects of a continuous policy monitoring including potential adjustments of the policies if necessary.

6 References

Banister, D. (1998, ed.). Transport Policy and the Environment. London and New York.

Banister, D., Hickman, R. (2005). VIBAT – Visioning and backcasting of UK transport policy. Three Reports for DfT, January, June and September.

Barbour, R. (2007). Doing Focus Groups. London.

Burgess, A., et al. (2005). TRANS-TOOLS (TOOLS for Transport forecasting ANd Scenario testing) Deliverable 1. Funded by 6th Framework RTD Programme. TNO Inro, Delft, Netherlands.

DLR and KIT (2010). OPTIC Deliverable 2: Inventory of tools and methods for early detection of adverse effects.

Grunwald, A. (2002): Technikfolgenabschätzung – Eine Einführung. Berlin.

iTREN (2009). iTREN - Integrated transport and energy baseline until 2030. Deliverable2, Project co-funded by the European Commission.

Kosow, H., Gaßner, R. (2008). Methoden der Zukunfts- und Szenarioanalyse. Überblick, Bewertung und Auswahlkriterien. WerkstattBericht Nr. 103. Berlin IZT - Institut für Zukunftsstudien und Technologiebewertung.

Morgan, D.L. (1997). Focus Groups as qualitative research. Qualitative Research Methods Series. A Sage University Paper. London.

Newton, S., Szimba, E., van Meijeren, J., Kraft, M., Schröder, S., Krail, M., van Herbruggen, B., Purwanto, J., Maurer, H., et al. (2009). Preparing TRANSTOOLS and TREMOVE for linkage with long-term energy and economic models. Deliverable D2 of iTREN-2030 (Integrated transport and energy baseline until 2030). Project co-funded by European Commission 6th RTD Programme. Zoetermeer, Netherlands.

Rich, J., Bröcker, J., Hansen, C.O., Korchenewych, A., Nielsen, O.A., Vuk, G. (2009). Report on Scenario, Traffic Forecast and Analysis of Traffic on the TEN-T, taking into consideration the external dimension of the union – TRANSTOOLS version 2; Model and Data Improvements, Funded by DG TREN, Copenhagen, Denmark.

Schippl, J., Leisner, I. (2009). Backcasting in Transport. Scenarios for the Future of European Long-distance Transport. Technikfolgenabschätzung – Theorie und Praxis Nr. 2, 18. Jg., September 2009. http://www.itas.fzk.de/eng/tatup/inhalt.htm

TSU Oxford et al. (2010). OPTIC Deliverable 1: Inventory of measures, typology of non-intentional effects and a framework for policy packaging.

Van Herbruggen, B. (2007). TREMOVE model description, Transport & Mobility Leuven, available at: http://tremove.org/documentation/TREMOVE_Short_Description.pdf