CPB Discussion Paper

No 27

February, 2004

How econometric models help policy makers
Theory and practice*

Henk Don

* This paper is based on a presentation for the Econometric Methodology Conference in Oslo, August 2003. Helpful comments by Frank den Butter, Peter Cornelisse, Peter van Els, Martin Fase and Bernt Stigum are gratefully acknowledged. The usual disclaimer applies.

CPB Netherlands Bureau for Economic Policy Analysis
Van Stolkweg 14
P.O. Box 80510
2508 GM The Hague, the Netherlands

Telephone +31 70 338 33 80
Telefax +31 70 338 33 50
Internet www.cpb.nl

ISBN 90-5833-158-X

Abstract in English

Frisch and Tinbergen founded the standard framework for finding the optimal economic policy by maximizing the welfare function under constraints supplied by the econometric model. Frisch worried about the reliability of the model and Tinbergen thought that it would be too difficult to specify the welfare function. Looking at current practice in Dutch policy making, both worries are relevant but the solutions proposed by the founders are not very helpful. Rather, the solution is found in applying an iterative trial-and-error procedure interfacing between the policy maker and the model-cum-expert system. The main contributions of the standard framework are its useful set of concepts, the famous order condition for a feasible solution, and the clear definition of role models for the two parties in the interaction.

Key words: macroeconomic models, policy choice, policy vigour, reliability, uncertainty, welfare function

Abstract in Dutch

Frisch en Tinbergen hebben de grondslag gelegd voor het standaard raamwerk om het optimale economische beleid af te leiden uit het maximeren van de welvaartsfunctie onder restricties die beschreven worden door het econometrische model. Frisch maakte zich zorgen over de betrouwbaarheid van het model en Tinbergen meende dat het te moeilijk zou zijn om de welvaartsfunctie te bepalen. In de hedendaagse praktijk van de beleidsvoorbereiding in Nederland blijken beide punten van zorg terecht, maar de oplossingen die door de grondleggers zijn voorgesteld, blijken niet erg behulpzaam. De oplossing wordt eerder gevonden in de herhaalde toepassing van een trial-and-error procedure die vorm geeft aan de communicatie tussen de beleidsmaker en het model met de bijbehorende expertise. De belangrijkste bijdragen van het standaard raamwerk zijn de nuttige concepten die het hanteert, de bekende orde-voorwaarde voor het bestaan van een oplossing, en de heldere definitie van rolmodellen voor de beleidsmaker en de expert.

Steekwoorden: macro-economische modellen, beleidskeuze, beleidskracht, betrouwbaarheid, onzekerheid, welvaartsfunctie

Een uitgebreide Nederlandstalige samenvatting is beschikbaar via www.cpb.nl.

Table of contents

Summary

According to the standard framework, optimal economic policy is found by maximizing the welfare function under constraints supplied by the econometric model. Frisch and Tinbergen, the two founders of this framework, both worried about its practical usefulness.

Frisch worried about the reliability of the econometric model for the purpose of giving policy advice. In particular, he felt that the behavioural equations could be unnecessarily restrictive for finding the best policy. If it is to be used for supporting policy decision making, a model should contain only "autonomous relations", defined as relations which can be expected to remain invariant under any change in the rest of the model. Because many behavioural equations cannot be expected to remain invariant under fundamental changes in institutions, these equations are unreliable if such institutional changes are among the policy decisions that one is willing to consider. This may be read as an early statement of the famous Lucas critique. Frisch proposed to solve the policy problem in two stages. In the selection phase, the welfare function is maximized under a subset of model equations, comprising only of definitions, accounting identities and technical limitations. This yields the best vector of targets in the set of physically possible results. Next, in the implementation phase, one tries to find the policy instruments that bring the economy as close as possible to the selected targets.

Tinbergen thought that in practice it would be too difficult to specify the welfare function which the optimal policy is supposed to maximize. This welfare function should reflect collective preferences over different states of the economy. It should not only depend on the individual utility functions of the citizens, but also on a certain way of combining and hence weighing these individual interests. As a short-cut solution, Tinbergen assumes that policymakers choose some fixed targets instead.

The formal treatment of uncertainty in the standard framework is largely due to Theil, who derived the famous certainty equivalence result: Under specific conditions, maximum expected welfare is obtained for the policy vector that maximizes welfare under certainty, substituting expected values for the uncertain elements. Unfortunately the conditions are unlikely to be met in macroeconomic applications. For policy analysis, the crucial uncertainties relate to the impact multipliers and the specification of the model. Brainard and Don have shown that an increase in uncertainty about the impact multipliers reduces the vigour of the optimal policy.

Looking at current practice in Dutch policy making, the worries of Frisch and Tinbergen both are relevant but the solutions proposed by them are not very helpful. Rather, the solution for both problems is found in applying an iterative trial-and-error procedure interfacing between the policy maker and the model-cum-expert system. Policy makers suggest a particular combination of instrument values, which is analyzed and translated into model input by the professional model user. The model results are supplemented by the expert and translated for the policy makers. The report is likely to lead to analytical questions and a discussion between policy

makers and experts, both on the quality of the analysis and on possible improvements of the policy proposals. A revised policy is then submitted to the experts for a fresh analysis.

The main contributions of the standard framework are its useful set of concepts, the famous order condition for a feasible solution, and the clear definition of role models for the two parties in the interaction. The distinction between target variables and instrument variables is crucial for any useful interaction between policy makers and experts. The same holds true for the concept of a welfare function describing the preferences of policy makers and an econometric model describing the relations between instruments and targets. The order condition, stating that a policy problem may only be solved if the number of instruments is at least as large as the number of targets, proves very useful in the interaction. And it is important that both parties in the interaction observe the responsibilities that are defined by the role models: the policy maker decides on the welfare function and ultimately chooses the instrument values that will be implemented; the expert decides on the model relations and determines the forecast for the target variables once the instrument values have been chosen.

Uncertainty in the baseline scenario is readily communicated to the policy makers and has been taken explicitly into account over the last decade in preparing Dutch fiscal policy. Perhaps the difficulties in assessing actual parameter uncertainty are an important obstacle in adjusting policy vigour to parameter certainty. In practice, policy vigour may well be determined by other considerations, ranging from primitive feelings that “something must be done” to sophisticated political strategies and the expected impact on public opinion.

The first priority for improving current practice in the Netherlands is extending the scope of the econometric models. In particular, we need sound estimates of the structural economic effects of government programs in education, infrastructure, health, etc. In addition, we should work on models suited to evaluate more policy options in changing institutions. The second priority is to find a proper way of determining and communicating the uncertainty inherent in any particular piece of policy analysis. There is no need for a new type of method for policy support.

1 Introduction

Frisch (1950) and Tinbergen (1952) established the theory of economic policy which still provides the standard textbook framework for discussing how econometric models are used to help policy makers.[1] However, Frisch and Tinbergen themselves noted that for a number of reasons the standard framework would be difficult to apply in practice. Frisch worried about the reliability of the econometric model for the purpose of giving policy advice. In particular, the behavioural equations could be unnecessarily restrictive for finding the best policy. Tinbergen felt that specifying a welfare function to be maximized is a difficult matter, which in practice would often be passed over by directly choosing a number of fixed targets.
This paper confronts the standard framework and these early worries with current practice in Dutch policy making. It shows that the worries of Frisch and Tinbergen were well founded, but that the solutions they suggested are not very helpful in practice. In practice, other ways have been found to solve their problems.

Current practice of using econometric models in Dutch policy making, as described below, shows many similarities with current practice elsewhere. Related accounts of policy support in practice may be found in Bray et al. (1995) and the collected papers edited by Britton (1989) and Den Butter and Morgan (2000).

[1] For Frisch's views, I should also refer to Frisch (1962) and the account given by Leif Johansen (1977). For Tinbergen, the prime source is Tinbergen (1952), where he is more explicit on the welfare function and its practical limitations than in most later work.

2 Theory

The framework for the theory of optimal economic policy was established by Frisch (1950) and Tinbergen (1952). The basic setup is quite simple. With y a vector of target variables, x a vector of policy instrument variables and z a vector of other (exogenous) variables, the econometric model is assumed to provide us with a function f describing what value of the target vector y results from a particular instrument vector x, under given values for the other variables z, i.e.

(1) $y = f(x, z)$.

The policy problem then is to maximize a welfare function (or preference function) W of the target variables subject to the restrictions given by the econometric model, i.e.

(2) $\max W(y)$ subject to $y = f(x, z)$ for given z

Both Frisch and Tinbergen noted that this simple setup would be difficult to apply in practice. Therefore they came up with adjustments designed to overcome practical problems in implementation. Interestingly, their worries went in different directions.

2.1 Reliability of behavioural equations

Frisch (1950) stresses that a model should only contain "autonomous relations" if it is to be used for supporting policy decisions. A relation is said to be autonomous if it can be expected to remain invariant under any change in the rest of the model (Frisch 1950, p. 485).[2] Because many behavioural equations cannot be expected to remain invariant under fundamental changes in institutions, they are not reliable if such institutional changes are among the policy decisions that one is willing to consider.[3] By imposing such unreliable equations as restrictions in the policy optimization, one could miss opportunities for a better result. Indeed, Johansen (1977) explains that the welfare function should be maximized under a subset of restrictions, comprising only of definitions, accounting identities and technical limitations. Thus one would select the best vector of targets in the set of "physically possible results". Only after this "selection phase" one should proceed to the "implementation phase", which consists of finding the set of policy instruments that can bring the economy as close as possible to the selected target vector. The latter phase calls for creative policy ideas, including institutional changes that may affect the behavioural relations. In this context, finding the best policy is characterized by interaction with the policy makers and the procedure is by trial and error rather than mechanical

[2] In his Econometrica study, Haavelmo (1944) explicitly acknowledges Frisch as the originator of this concept of autonomy.
[3] Note that a model containing only autonomous relations is immune to the Lucas critique. We return to this issue in section 3.1 below.

optimization. The model relations are accepted only if they do not unnecessarily preclude reaching the best possible result.

In contrast, Tinbergen (1952) relies on all model equations in finding the optimal instrument vector. But he states clearly that his method is restricted to the determination of "quantitative policies". This rules out "qualitative policies" and "reforms", which affect the institutional structure and hence the model relations. The study of such policies is more difficult and explicitly outside the scope of his 1952 contribution. In Tinbergen (1956) he devotes one chapters to qualitative policies, defined as changes in structure "within given foundations". Thus chapter 5 studies built-in stabilizers, changes in pricing and taxation schemes, monopolies, decentralization or centralization in administration and the appraisal of investment projects. Chapter 6 discusses several types of reforms, defined as "changes in foundations". Both types of policy questions can be handled in the standard framework, provided the model is sufficiently broad to encompass the relevant policy options. In practice, it requires dedicated models to study various types of institutional change. As Tinbergen (1952, p. 72) noted, such problems meet with great difficulties, mainly because "our empirical quantitative knowledge of human behaviour under different structural conditions is so restricted" and he warns against speculative and biased solutions.

2.2 The unknown welfare function

Tinbergen expects that the standard framework cannot be applied because in practice the welfare function W is unknown. This function reflects collective preferences over different states of the economy. It should not only depend on the individual utility functions of the citizens, but also on a certain way of combining and hence weighing these individual interests. Right at the start Tinbergen states that the specification ("fixation") of the welfare function is "a difficult matter; generally it will not be considered consciously but intuitively by those responsible for the policy. (…) In practice the stage of fixing (W) and trying to maximise it will often be passed over and the targets y chosen directly" (Tinbergen 1952, p. 2-3). So as short-cut solution, he proposes to choose a fixed target vector y^* and replace (2) by

(3) solve $y^* = f(x, z)$ for x at given z

This will yield the corresponding instrument values. Tinbergen anticipates that the result will not prove satisfactory, because it will violate either some technical constraints (nonnegative prices, technical possibilities) or some social or political constraints. Hence a critical discussion will give rise to various boundary conditions that must be observed and Tinbergen then goes on to discuss their inclusion in the procedure to find the proper instrument vector. If the solution to (3) violates technical constraints, clearly the model (1) is incomplete or only locally valid. If the solution violates political constraints, the replacement of the welfare function W by a fixed set

of targets y^* was inadequate. In both cases, what follows is a trial-and-error procedure which is intended to successively incorporate all relevant constraints. Of course, the policy maker could also suggest a different target vector, but apparently that option is less interesting from Tinbergen's point of view.

Replacing the maximization of $W(y)$ by fixing the target vector y^* turns the policy problem (2) into a set of equations (3). Of course existence of a solution (for arbitrary y)[4] requires the famous order condition[5], i.e. the number of instrument variables should be at least as large as the number of target variables, i.e. dim $x \geq$ dim y.

Frisch is well aware that the specification of the welfare function is no easy matter. Indeed he has worked a lot on both the theoretical and practical aspects and he reached a rather definite opinion, based not least on extensive interviews with leading policy makers: "I am convinced that the preference function problem for an economy at large can be solved when it is approached in an intelligent and cautious way" (Frisch 1962, p. 255). Yet, it cannot be formulated in one stroke, but should be found "through a series of attempts based on continuous cooperation between the responsible authorities and the analytical experts" (ibidem).

2.3 Uncertainty

The formal treatment of uncertainty in the present context was developed by Theil (1954, 1958). He derived the famous certainty equivalence result, stating that under certain conditions, maximum expected welfare is obtained for the policy vector that maximizes welfare under certainty, substituting expected values for the uncertain elements.

One important condition for certainty equivalence is that the welfare function is quadratic in the target variables. This condition is unlikely to be met in macroeconomics, even symmetry is unlikely to hold as policy makers tend to get into deeper trouble in case of setbacks than in case of windfalls. This relates to the time lags involved in determining the relevant economic variables and in adjusting policy decisions to new information. Setbacks are more difficult to handle than windfalls. As a result, certainty equivalence is not very helpful in macroeconomic policy practice;[6] see also the discussion in Don (2001a).

In addition, the certainty equivalence conditions exclude uncertainty in the (multiplicative) model coefficients which relate the targets to the instruments. Even in a static linear-quadratic optimization model, uncertainty about the impact multipliers does affect optimal policy. Here the seminal result is that of Brainard (1967), who showed that in the single target, single

[4] If the order condition is not satisfied, the set of reachable target vectors is only a subset of the space of possible target vectors. In that case a solution exists only for target vectors in the reachable subset.

[5] The order condition is necessary but not sufficient. To ensure sufficiency, it should be supplemented by a rank condition.

[6] The parametric certainty equivalence procedures proposed by Johansen (1980) allow for particular forms of risk aversion and asymmetry. They lead to the maximization of some linear combination of expected welfare and its variance, which cannot handle the asymmetry indicated in the text.

instrument case an increase in the uncertainty about the impact multiplier reduces the vigour of the optimal policy. For a high level of uncertainty, this confirms the old adage “when in doubt, abstain”. Don (1983) provided an unconditional generalization to the multidimensional case.

Frisch probably considered these uncertainties as a minor issue, compared to the major issue of model reliability discussed above. Tinbergen (1954) did worry about uncertainty in the impact multipliers. He explained that information from different sources helped to reduce parameter uncertainty. In addition he suggested to perform sensitivity analyses with respect to those parameters that were considered to be less certain. Indeed, he did exactly that in Tinbergen (1952, p.60-61) for a small macroeconomic model. Targeting production and the balance of payments with government expenditures and the exchange rate, he studied the sensitivity of the optimal policy with respect to four model parameters: the marginal propensity to spend, the price elasticities of imports and exports, and the marginal wage quota (share of labour costs).

3 Practice

How do econometric models in fact help policy makers? Does the simple setup as described by (1) prove useful in practice? How important are Frisch's worries about the reliability of the model and Tinbergen's worries about knowledge of the welfare function, and how workable are their suggestions to cope with these practical problems?

The simple setup works fine for relatively small and simple models. There it helps exploring the trade-offs that policy makers face. In a multi-player setting the simple setup provides a good framework for studying the characteristics of the policy game. Players may be different policy-making bodies in a single country (government, central bank, trade unions) or different national governments in a multi-country setting.

However, such studies are mainly used to get a general idea of the type of dilemmas and choices that the policy makers face. While small models can capture the characteristics of a particular policy problem, they are necessarily incomplete and often unsuitable for guiding actual policy choices.[7] Real policy making tends to draw on larger and more sophisticated models, that more accurately describe the institutional setting and the range of specific policy instruments. Such "large" models are required to cope with the complex interdependencies in the real economy in a way that is accountable, traceable and consistent .[8] So let us see how the ideas of Frisch and Tinbergen apply to a realistic policy problem in the context of a "large" empirical econometric model.

3.1 Usefulness and reliability of large models

Models are simplifications of reality. As Tinbergen wrote in 1936:
"To get a clear view, things must be stylized. The many phenomena must be grouped in such a way that the picture becomes clear, yet without losing its characteristic traits. (…) Some have made stylized pictures that were unwieldy. Some have made stylized pictures that were unrealistic. But things must be stylized. The alternative is barrenness".
(from the Dutch original in Tinbergen (1936);[9] translation by the present author)

Tinbergen writes about *stylized* pictures rather than *simplified* pictures. This indicates that the simplification of reality should be well structured: it should obey some set of rules, constituting a scholarly style. Rules are readily offered by economic theory, logic and the relevant

[7] Unlike one or two decades ago, it now seems to be taken for granted that real world decision support for macroeconomic policy requires the use of large models, *e.g.* Pagan (2003) and Sims (2002).
[8] To illustrate, the coalition agreement reached by three Dutch political parties in May 2003 counted about 60 different economic policy measures (CPB, 2003b, appendix A). In an intermediate analysis, these were translated into changes in about 30 exogenous variables of the JADE model, see CPB (2003a).
[9] The full text of this paper is available in English under the name "An economic policy for 1936", in Klaassen, Koyck and Witteveen (eds., 1959), pp. 37-84. There the quoted phrases are found on p. 41.

accounting system. This is how a model provides consistency with economic theory. In addition, using a model provides consistency over time. Taking possible changes in the model into account, the present model analysis is consistent with earlier analyses.

The stylized picture must be manageable and realistic. Of course, the size of models that can be managed nowadays is much larger than in 1936 or 1952. Indeed models have grown larger to enhance their realism. But still they are simplifications, leaving out elements that are considered to be inessential for the problems at hand. There are no models suitable for all economic policy problems. Different models are built for different purposes. For empirical macroeconomic models, the restriction on size nowadays does not come from limitations in computing power, but from what can be traced and understood by the user and the client. Black boxes are useless. If the model user cannot explain the results, his clients will not trust the analysis. Indeed, the model should offer a clear view on the relevant phenomena and their interaction. The user must be accountable for the model analysis to his colleagues and clients. He must be able to trace and understand the model results. For more complex models, this requires a knowledgeable user and sophisticated communication with the clients.

How about reliability? Being only a stylized picture, the model is always incomplete. Phenomena that were considered inessential at the model building stage, may prove to be important in a particular application. Hence the professional user must be aware of the limitations of the model and be ready to adjust the model or supplement the model analysis to cope with relevant issues outside the scope of the standard model version.[10] Often the incompleteness also means the omission of alternative institutional arrangements. Unless the model was built to help with problems of institutional design, it is likely to take the current institutional setup as given. This common situation rules out using the model for analysing what Tinbergen called qualitative policies and reforms.

Wherever the model goes beyond accounting identities, uncertainties creep in. Behavioural relations are not known with precision, but have been estimated and therefore are subject to uncertainty. Additional uncertainty relates to the possible instability of the behavioural relations over time. Ideally, the behaviour of an agent in the model is described as the result of the restricted optimization problem that he faces. Thus behaviour can be derived from preferences, institutions and technology, in a way which makes the behavioural relations invariant under changes in the rest of the model. This answers Frisch's call for autonomous relations, discussed above. It also answers the famous Lucas critique: Lucas (1976) argued that often some model parameters are unlikely to be invariant under policy change, and hence the model cannot be used to assess the effects of a policy change. In the ideal model, any change in policy regime will affect the restricted optimization problem only through changes in the exogenous variables and can be handled correctly. Even then, we have no assurance that preferences and technology

[10] See Turner et al. (1989) for an insightful account of how macroeconometric models are used in practice to evaluate different types of policy proposals, creatively handling various difficulties that one encounters along the way.

are sufficiently stable over time to allow us to use the model for establishing an optimal policy for the future. And, of course, in practice the stylized pictures that we use cannot live up to the ideal. Hence our models are unreliable for a number of reasons, as the behavioural relations are based on imprecise estimates and may suffer from various types of instability.

Another imperfection in actual models is particularly relevant to their use in policy optimization. Many model relations can claim only local validity, because they have been derived as local approximations to more complex and usually unknown global relations. In an optimization exercise, technically the optimum may well lie outside the local area of reliability.[11] For example, in an optimization exercise I did almost two decades ago with a linearized version of the CPB macro model of that time, the technical optimum implied an abolition of social security premiums paid by employers and a raise in the VAT by the same amount (Don, 1986a). This was the result of a small incongruence in tax shifting parameters in the wage equation, which may well hold true for small changes but is unlikely to remain valid for the large shift suggested by the technical optimum. This is the type of misleading result that Tinbergen hoped to exclude by adding proper boundary conditions.

While a large empirical macroeconomic model is highly useful for coping with complexity and for ensuring consistency and accountability, it is necessarily incomplete, uncertain and often only locally valid. The model is best used as a valuable partner in the discussion. This partner has a lot of patience, an excellent memory and superior computing skills. Surprises in model results can mean a lesson for the model users, but more often indicate an error in the model code or input data. To cope with its limitations, the professional user must be creative, focussed and future-oriented. One should not have blind faith in the model results, but always question their validity for the problem at hand.

Technical optimization of large empirical models tends to yield useless results. Hence, in practice, policy choice proceeds by trial and error.[12] Policy makers suggest a particular combination of instrument values, which is analysed and translated into model input by the professional model user. The model results are supplemented by the expert and translated for the policy makers. The report is likely to lead to analytical questions and a discussion between policy makers and experts, both on the quality of the analysis and on possible improvements of the policy proposals. A revised policy is then submitted to the experts for a fresh analysis.

So Frisch was certainly right in worrying about the unreliability of the model relations. And yes, part of the answer is performing a trial-and-error procedure. But stripping the model of all

[11] Or, as Bray et al. (1995, p. 997) put it: “In practice, what tends to happen with an empirical macromodel is that there will be some odd quirk in the model, an odd nonlinearity, a corner solution or even an extreme assumption such as rational expectations which the optimal policy rule is able to exploit.”

[12] According to Johansen (1977, p. 153), a book by Jöhr and Singer (1955) refers to intuition and judgement combined with a trial and error procedure as the superior method which solves the problem in practice, in spite of the fact that a complete formalization would involve almost insurmountable difficulties.

its behavioural relations does not really help. A study by the Dutch Scientific Council for Government Policy followed Frisch's suggestion to select target values in an optimization exercise ignoring the behavioural relations (WRR, 1987). This study showed that technically there was scope for higher growth, lower unemployment and less environmental damage than the CPB analyses of that time tended to suggest. Indeed, just observing technical input-output relations and not worrying about the price elasticity or composition of final demand, one could come up with a specialization pattern yielding high growth and low environmental damage. However, after this selection phase the Council did not really succeed in finding a set of policies that would make it happen (the implementation phase). Barring a full fledged Soviet type planning system, it did not give a clue on the type of policy, if any, that would be needed to steer the economy towards the technical optimum.[13] CPB's empirical behavioural relations may well have been too restrictive for finding the best possible policy. Creative ideas on how to overcome particular restrictions and expand the set of reachable targets cannot be delivered by the model, but must come from policy makers and experts. An optimization with fewer behavioural restrictions may give inspiration for generating creative ideas, but ignoring all of them does not really help.

As Klein observed, "(e)conomic policy is as much directed towards changing the constraint system as it is towards searching for exogenous variable configurations within a given control system" (Klein 1983, p. 160). It leads him to commend "the econometrician's traditional way of implementing system results for policy purposes. Instead of looking for the optimum, he studies various simulations with changing instruments, changing parameters, and changing equation specifications until he finds improved or possibly target value solutions. This is a more flexible procedure than formal search for the optimum in a given constraint system" (Klein 1983, p.160).

3.2 Getting to know the welfare function

If only for lack of technical skills, policy makers are unlikely to provide the expert with a well defined welfare function. In practice, preferences are formed and specified in the course of a trial-and-error procedure like the one I described in the previous section. When confronted with the likely outcome of a concrete policy proposal, policy makers will not only look for a better set of instrument values but also come up with additional targets or conditions that must be observed. The expert also has a role to play here. He should make sure that the report on the likely outcome of a concrete policy proposal is complete, i.e. that it contains information on all variables that might be relevant to the policy maker. For instance, if an income distribution target is proxied by a particular ratio of quantiles, any special impact on other quantiles should be reported so that it may be added to the list of targets if the policy maker cares about it. "No pain, no gain" is a useful rule of thumb: if a revised policy proposal yields an improvement on

[13] For critical reviews of the WRR study, see Don (1986b) and Fase (1987).

some target, a price is likely to be paid on some other target. If this price is not immediately identified, the expert should go and look for it, possibly by expanding the list of targets.

The trial-and-error process constitutes an interaction between the policy maker and the model-cum-expert system. This interaction is crucial in getting messages across in two directions. It is how the policy maker learns about trade-offs between different targets and the feasibility of particular combinations of target values. It is how the expert learns about the relevant list of target variables and the weights attached to them. In the course of the exchange, different types of discussions will come up. The policy maker will question the limitations of the model and ask for model validation, model extensions or supplementary analyses. The expert will elicit the exact political targets and weights. If the process is cooperative and open-minded, together they will engage in a creative process to find new types of instruments that enlarge the set of feasible targets by moving the various trade-offs away from the origin. Such new instruments may lie outside the immediate scope of the model, and include what Tinbergen called “qualitative policies”.

So Tinbergen was certainly right in worrying about the identification of the welfare function. And yes, part of the answer is performing a trial-and-error procedure. But choosing fixed targets for optimization and subsequently adding constraints does not really help. Of course, an optimization exercise with fixed targets may help to illustrate what type of solution could be available. Often adding constraints will indeed be necessary to avoid useless results outside the local area of model validity discussed above. But getting to know the welfare function requires some way of tracing the relevant trade-offs between targets and getting feedback from policy makers on the positions they prefer. This does not require technical optimization but is readily accomplished by the trial-and-error procedure outlined above.

Interestingly, there are ways to recover the welfare function from the policies that finally are advocated, or indeed decided. Aslaksen and Bjerkholt (1985) studied the preferences implicit in long-term strategies for depletion of Norwegian oil reserves. Brandsma et al. (1988) studied the preferences implicit in an illustrative macroeconomic policy package designed by CPB (1986). In order to get plausible results both studies had to account for risk aversion related to the uncertainty in the baseline scenario.

3.3 Handling uncertainty

The baseline scenario usually is a macroeconomic forecast based on the assumption that current policies are unchanged. It serves as a reference scenario for discussing the effects of alternative policy options. Policy makers readily understand that the baseline scenario is characterized by major uncertainties. Their asymmetric loss function, discussed in section 2.3 above, may lead them to consider a cautious scenario, or more precisely a scenario which holds a stronger policy

challenge, rather than a central forecast. This has indeed been standard practice since 1994 for preparing budgetary policies in the Netherlands. Yet it has not been applied to other policy fields, if this would mean using different scenarios for different policy questions; see Don (2001a) for a discussion. While this particular usage of scenarios may be typical for the Dutch situation,[14] one may also achieve the same result in less formal ways.

While forecast uncertainty in macroeconomics is dominated by the uncertainty in exogenous variables and error terms in behavioural equations (Don (1994)), the major sources for uncertainty about estimated policy effects are the model specification and the model parameters. If uncertainty about policy effects is derived only from the estimated standard errors of the parameter estimations, the results will be misleading. On the one hand, the estimated standard errors understate the true uncertainty because they are only valid on the assumption that the model specification is correct. As Wallis (1993, p. 126) put it, "no measure of uncertainty that surrounds the choice of (the underlying view of the world) is available." On the other hand, the estimated standard errors may well overestimate the true uncertainty because they can only reflect the inaccuracy of the parameters in the context of the sample used for estimating them, whereas in most cases other sources of knowledge have gone into the choice of the parameter values used in the policy model. As Tinbergen (1954, p. 242) observed, "the uncertainty margins obtained (from statistical reliability measures) are often wider than the uncertainty margins that our economic judgement would give us".

In principle, one could set up a Bayesian analysis that takes proper account of the various sources of uncertainty. However, this is highly impractical, especially in the context of a large real world policy model.[15] In practice, one hardly gets beyond a well chosen set of sensitivity analyses, following the example of Tinbergen (1952). This is far from satisfactory, indeed Turner et al. (1989) warn us that the interpretation of such analyses is essentially subjective. Perhaps the difficulties in assessing actual parameter uncertainty are an important obstacle in adjusting policy vigour to parameter certainty, in line with the analysis of Brainard (1967). In fact I think policy vigour tends to be determined by other considerations, ranging from primitive feelings that "something must be done" to sophisticated political strategies and the expected impact on public opinion.

[14] Dutch practice has even gained in sophistication since its description in Don (2001a). Don (2001b, section 7) explains how a probability statement is attached to the cautious scenario.

[15] Sims (2002) appears more optimistic on this issue, but notes that in practice so far one has used other methods, largely relying on subjective judgemental probability distributions.

4 What makes the theory useful

So both Frisch and Tinbergen were right in their worries about the practical usefulness of their common framework for the theory of optimal economic policy, as summarized in (1). Realistic policy problems involve “large” models containing uncertain relations which may be too restrictive and often can claim only local validity. The welfare function describing the preferences of the policy maker is unknown and hard to identify. The suggestions Frisch and Tinbergen gave to cope with these problems tend to be of little help. Rather, an open interaction between policy maker and the model-cum-expert system can solve the policy problem using a straightforward trial-and-error procedure.

What is left to make the theory useful? First, a crucial contribution is the set of *concepts* that the theory introduced. The distinction between target variables and instrument variables is central to any useful interaction between policy maker and expert. The same holds true for the concept of a welfare function describing the preferences of the policy makers and an econometric model describing the relations between instruments and targets. Second, the celebrated *order condition* provides a valuable intuition. That a policy problem may only be solved if the number of instruments is at least as large as the number of targets, is a condition which is easily communicated to the policy maker and proves very useful in the interaction. Third, the framework defines important *role models* for the two parties in the interaction.[16] The policy maker decides on political preferences (the welfare function) and chooses the policy that will be implemented (instrument values). The expert decides on the model relations (including additional restrictions or supplementary analyses) and determines the forecast for all other relevant variables, including the target values (conditional on the instrument values chosen by the policy maker). While each party can question the decisions of the other, these role models define the ultimate responsibilities that should be observed.

4.1 Illustrations from recent experience in the Netherlands

The role models are essential in the process of assessing the economic effects of different election platforms, which has been performed in the Netherlands in the run-up to general elections since 1986. For an in-depth discussion of the merits and drawbacks of this process, see the various contributions in the book edited by Graafland and Ros (2003).

In the run-up to the 2002 elections no less than seven political parties asked to be part of the exercise, indeed comprising all parties then represented in parliament. Each party was quite happy with the results of the assessments. Apparently, the differences in estimated effects between the party platforms reflected the differences in political preferences (welfare functions) between the parties. Table 1 offers a good illustration. Because of the workload involved and the time constraint, the trial-and-error procedure could not have many steps. On the basis of the

[16] Cf. Johansen (1977), Section 2.5 (pp. 104-110).

published policy platforms, CPB asked parties to elaborate and specify their policy proposals. This requires the constructive cooperation of party officials. After a first trial assessment, the party was given the opportunity to adjust its policy package so as to obtain a better result. After this second trial, a third trial was allowed only to correct errors and misunderstandings.[17]

Table 4.1 Selection of estimated effects of two policy platforms

	Progressive environmentalists (GreenLeft)	Conservative liberals (VVD)
Structural growth of GDP	–	+
Equitable income distribution	+	–
Public spending	+	–
Structural budget balance	–	+
Environment	+	–

Source: CPB (2002).

The role models require a clear demarcation of available instruments, because the policy maker decides on the instrument values and the expert takes responsibility for forecasting the other variables. In the current Dutch institutional context, this means that the wage rate cannot be considered an instrument, unless the policy maker declares his intention to change the relevant laws and take control of wages. It also implies that policy variables that are controlled either by international bodies or by local authorities cannot be considered as instruments for the purpose of assessing party platforms for national parliamentary elections. Part of the value of the exercise lies in the fact that the proposals for economic policy are made concrete and mutually comparable by means of a common baseline scenario and a standard reporting framework. The cabinet that was formed after the May 2002 elections broke up in October of the same year. New elections were held in January 2003, but because of time constraints CPB could not do a proper assessment of party platforms on this occasion. Thus we had a natural experiment on the value of such an assessment. The lack of it led to quite some confusion in the public debate on what the economic policies advocated by different parties actually implied. Only a week or so before the elections it transpired that the budget cuts advertised by one party were obtained by adding up the cuts in four consecutive years, while others only listed the impulse in the fourth year. In addition to such differences in definitions used, various claims about the effects on target variables such as employment and the government budget went unchecked. This left substantial scope for the delivery of hollow promises and the omission of unfavourable effects (Van Wijnbergen and Beetsma, 2003).

Because no single party has a majority in parliament, after the elections two or more parties start negotiations to form a coalition cabinet. Both in 2002 and in 2003, the negotiating parties

[17] For a more detailed account of the procedure and the instruments used, see Don (2003).

called on CPB to support the search for an economic policy program that would best serve their preferences. Again, this took the form of a trial-and-error procedure. The negotiating parties submitted a trial policy package to CPB, which gave an assessment of the likely economic effects. More often than not, the results would not quite fit the preferences of the parties, thus prompting them to submit a new trial package, and so on. After the January 2003 elections, the first round of negotiations was between the two parties that had obtained the best results in the elections, the Christian Democrats (CDA) and the Social Democrats (PvdA). In the course of this process, it became clear that the targets they had set themselves could not be reached: it proved impossible to achieve a balanced budget by 2007 in the cautious scenario for the medium term without harming structural growth and employment and without violating various other political side conditions on the admissibility of instruments and on the income distribution. Efforts to find a feasible compromise failed and almost three months after the elections the negotiations between these two parties broke up. Next it took about four weeks for CDA to reach an agreement with the Conservative Liberals (VVD) and the Social Liberals (D66).

What kind of messages did the model-cum-expert system send to the negotiating parties? The trial-and-error procedure revealed several trade-offs that the policy makers had to face. Achieving budget balance by 2007 in the cautious baseline scenario required substantial cuts in government expenditures and/or increases in tax rates. On average, these measures suffer some 40% budgetary leakage because of their contractionary impact on the domestic economy. However, some measures suffer no leakage (*e.g.* reducing foreign aid) while other measures suffer high leakage (*e.g.* reducing government employment). Most expenditure cuts have only temporary contractionary effects because they do not hurt potential GDP growth, which is determined by structural productivity growth, labour supply and equilibrium unemployment. Indeed, some expenditure cuts, in particular in social security, may promote structural growth by stimulating labour supply and reducing equilibrium unemployment. On the other hand, most tax increases hurt structural growth by reducing labour supply and raising equilibrium unemployment. Again, some tax increases are less harmful than others, depending on how they affect total labour costs directly and through tax shifting. Unfortunately, some of the less harmful options for raising taxes met with a political taboo, in particular those related to the fiscal treatment of owner occupied houses.

Of course, all the macroeconomic considerations have to be weighed against the more microeconomic impact of various types of government expenditures and tax changes. Also, the beneficial long term macroeconomic effects of government spending on education, infrastructure, health care, public safety and social security should not be ignored, even if such effects still elude quantitative assessment.

These messages to the policy makers may all sound fairly straightforward, yet several elements were challenged in the public debate. More importantly, hardly anyone blamed CPB for the political choices made by the negotiating parties. The responsibilities of the politicians

were clearly separated from those of the model-cum-expert system, in line with the role models laid down by Frisch and Tinbergen. Indeed, the opposition parties used the CPB assessment of the final coalition agreement to attack the political preferences of the new coalition

4.2 Priorities for improvement in the Dutch case

In the Dutch situation, the first priority for improving policy support certainly is to extend the scope of the analysis. Three types of expansion are in order. First, the direct temporary Keynesian effects of government spending should be supplemented with the indirect structural effects of the programs that the money is spent on, whether it is education, infrastructure, health care, or any other area of public concern. This is no easy task and one should not expect easy answers. The long term economic effects of any government program depend not only on the amount of money spent but even more on the quality of the program and the incentive structures it creates. Second, the border between quantitative policies and qualitative policies (and reforms) as defined by Tinbergen must be pushed further out. Since his days, we have made quite some headway by developing applied general equilibrium models that can assess institutional reforms in various fields, ranging from international trade to social security. Still, more remains to be done. But we should also remember Tinbergen's warnings against speculation and bias where sound empirical evidence is lacking. Indeed, some types of policy cannot be analysed with quantitative models, *e.g.* "where the new policy is diffuse, with possibly wide-ranging macroeconomic consequences, being intended to promote broad structural change that cannot be quantified (…): (some) industrial policy proposals fall into this category." (Turner et al. (1989), p. 104). One may want to try and use more qualitative cost-benefit-type of analyses to study such reforms, *e.g.* new regulations for competitive market structures. Third, there are always some specific weak spots in the current system of models that require attention. Looking at our recent experience, we lacked good estimates of cross-border tax evasion and asset price formation in relation to capital income taxes and asset taxes.

As a second priority, I think we should find a proper way to determine and communicate the uncertainty that is inevitably attached to our policy analyses. While communicating uncertainty to policy makers should not be taken for granted, determining the relevant uncertainty comes first.[18] This is much more difficult for policy analyses than it is for forecasts, if only because in policy analysis we have no equivalent of *ex post* measured forecast errors. Several practical problems were mentioned in section 3.3 above. Graafland (2003) suggests some steps forward, but these do not address all difficulties and tend to be rather subjective. Still, an assessment of the uncertainty attached to our policy analyses would prove very valuable, both for informing our clients and for deriving priorities for empirical research.

[18] In Don (2001a) I discussed the determination and communication of uncertainty in macroeconomic forecasts.

5 Conclusions

Any real world macroeconomic model used to support policy choice can claim only limited reliability. It is necessarily incomplete, sometimes ill suited for the problem at hand and often only locally valid. By implication, a mechanical optimization procedure to find the best economic policy is likely to be misleading, if not entirely out-of-order. Yet the model can be very helpful in assessing the likely consequences of different policy options, provided it is used by knowledgeable experts, who are well aware of the limitations of the model at hand and can contribute creative ideas on handling any shortcomings of the model in a particular policy analysis. The interaction between the policy maker and the model-cum-expert system then takes the form of an iterative trial-and-error procedure. This procedure can also cope quite naturally with the lack of knowledge on the welfare function. Policy makers submit policy proposals they hope will give them a better result. Helped by his model(s), the expert returns an analysis of likely effects. Next there is likely to be a serious discussion on both the merits of the economic analysis and the options for improving the policy from the perspective of the policy makers. There is no need for a better type of method here.

In the Dutch case, the first priority for improvement is in extending the scope of the models. In particular we need sound estimates of the structural economic effects of government programs in education, infrastructure, health, etc. In addition, we should work on models suited to evaluate more policy options in changing institutions. A second priority is to get a better grip on the uncertainty inherent in any particular piece of policy analysis.

References

Aslaksen, I. and O. Bjerkholt, 1985, 'Certainty Equivalence Procedures in Decision-Making under Uncertainty: an Empirical Application,' in: F.R. Førsund, M. Hoel and S. Longva (eds.), *Production, Multi-Sectoral Growth and Planning*, Amsterdam, North-Holland.

Brainard, W., 1967, 'Uncertainty and the Effectiveness of Policy,' *American Economic Review*, 57, pp. 411-425.

Brandsma, A.S., G.M.M. Gelauff, B. Hanzon and A.M.A. Schrijver, 1988, 'Retracing the Preferences behind Macroeconomic Policy: the Dutch Experience,' *De Economist*, 136, pp. 468-490.

Bray, J., S. Hall, A. Kuleshov, J. Nixon and P. Westaway, 1995, 'The Interfaces between Policy Makers, Markets and Modellers in the Design of Economic Policy: an Intermodel Comparison,' *The Economic Journal*, 105, pp. 989-1000.

Britton, A., ed., 1989, *Policymaking with Macroeconomic Models*, Aldershot, Gower.

CPB, 1986, *Centraal Economisch Plan 1986*, The Hague, Sdu Publishers.

CPB, 2002, *Keuzes in Kaart 2003-2006 – Economische Effecten van Acht Verkiezingsprogramma's*, The Hague, CPB; summarized in English in: Charting Choices 2003-2006; Economic Effects of Eight Election Platforms, CPB Document 19, The Hague.

CPB, 2003a, JADE, a Model for the Joint Analysis of Dynamics and Equilibrium, CPB Document 30, The Hague.

CPB, 2003b, Economische Gevolgen van het Strategisch Akkoord 2003-2006, CPB Document 22, The Hague.

Den Butter, F.A.G. and M.S. Morgan, eds., 2000, *Empirical Models and Policy-making: Interaction and Institutions*, London, Routledge.

Don, F.J.H., 1983, 'Uncertainty and the Vigour of Policy – a Stronger Result,' *Journal of Economic Dynamics and Control*, 6, pp. 187-191.

Don, F.J.H., 1986a, internal CPB memos dated February 20, 24 and 27.

Don, F.J.H., 1986b, ‘Modelkeuze en Modelgebruik,’ *Economisch Statistische Berichten*, 71, pp. 934-936.

Don, F.J.H., 1994, ‘Forecast Uncertainty in Economics,’ in: J. Grasman and G. van Straten (eds.), *Predictability and Nonlinear Modelling in Natural Sciences and Economics*, Dordrecht, Kluwer, pp. 569-581.

Don, F.J.H., 2001a, ‘Forecasting in Macroeconomics: a Practitioner's View,’ *De Economist*, 149, pp. 155-175.

Don, F.J.H., 2001b, ‘The Growth Potential of the Dutch Economy on Medium Term’, in: *Reflections on Economics and Econometrics – Essays in Honour of Martin M.G. Fase*, Amsterdam, De Nederlandsche Bank, pp.101-111.

Don, F.J.H., 2003, ‘Economic Analysis of Election Programmes: What, How and Why?,’ in: J.J. Graafland and A. Ros (eds.), *Economic Assessment of Election Programmes – does it make sense?*, Dordrecht, Kluwer, pp. 21-30.

Fase, M.M.G., 1987, ‘Meer Groei met Minder Vervuiling,’ *Economisch Statistische Berichten*, 72, pp. 852-858.

Frisch, R., 1950, ‘L’Emploi des Modèles pour l’Élaboration d’une Politique Économique Rationelle,’ *Revue d’Économie Politique*, 60, pp. 474-498; 601-619; 620-634.

Frisch, R., 1962, ‘Preface to the Oslo Channel Model: a Survey of Types of Economic Forecasting and Programming,’ in: R.C. Geary (ed.), *Europe’s Future in Figures*, Amsterdam, North-Holland.

Graafland, J.J., 2003, ‘Balancing Information and Uncertainty,’ in: J.J. Graafland and A. Ros (eds.), *Economic Assessment of Election Programmes – Does It Make Sense?*, Dordrecht, Kluwer, pp. 61-79.

Graafland, J.J. and A. Ros, eds., 2003, *Economic Assessment of Election Programmes – Does It Make Sense?*, Dordrceht, Kluwer.

Haavelmo, T., 1944, ‘The Probability Approach in Econometrics,’ *Econometrica*, 12, Supplement, pp. iii-vi + 1-115.

Johansen, L., 1977, *Lectures on Macroeconomic Planning 1*: General Aspects, Amsterdam, North-Holland.

Johansen, L., 1980, 'Parametric Certainty Equivalence Procedures in Decision-Making under Uncertainty,' *Zeitschrift für Nationalökonomie*, 40, pp. 257-279.

Jöhr, W.A. and. H.W. Singer, 1955, *The Rôle of the Economist as Official Adviser*, London, Allen and Unwin.

Klaassen, L.H., L.M. Koyck and H.J. Witteveen, eds., 1959, *Jan Tinbergen Selected Papers*, Amsterdam, North-Holland.

Klein, L.R., 1983, *Lectures in Econometrics*, Amsterdam, North-Holland.

Lucas, R.E.,1976, 'Econometric Policy Evaluation: a Critique,' in: K. Brunner and A.H. Meltzer (eds.), *The Phillips Curve and Labor Markets*, Amsterdam, North-Holland, pp. 19-46.

Pagan, A.R.,2003, 'Report on Modelling and Forecasting at the Bank of England,' *Bank of England Quarterly Bulletin*, Spring, pp. 1-29.

Sims, C.A., 2002, 'The Role of Models and Probabilities in the Monetary Policy Process,' paper prepared for the Fall meeting of the Brookings Panel on Economic Activity.

Theil, H., 1954, 'Econometric Models and Welfare Maximisation,' *Weltwirtschaftliches Archiv*, 41, pp. 60-83.

Theil, H., 1958, *Economic Forecasts and Policy*, Amsterdam, North-Holland.

Tinbergen, J., 1936, 'Kan hier te lande, al dan niet na Overheidsingrijpen, een verbetering van de binnenlandse conjunctuur intreden, ook zonder verbetering van onze exportpositie?,' *Praeadviezen voor de Vereeniging voor de Staathuishoudkunde en de Statistiek*, The Hague, pp. 62-108.

Tinbergen, J., 1952, *On the Theory of Economic Policy*, Amsterdam, North-Holland.

Tinbergen, J., 1954, 'Over de Theorie der Economische Politiek, *De Economist*, 102, pp. 241-249.

Tinbergen, J., 1956, *Economic Policy: Principles and Design*, Amsterdam, North-Holland.

Turner, D.S., K.F. Wallis and J.D. Whitley, 1989, 'Using Macroeconometric Models to Evaluate Policy Proposals,' in: A. Britton (ed.), *Policymaking with Macroeconometric Models*, Aldershot, Gower.

Van Wijnbergen, S. and R. Beetsma, 2003, 'Partijen Bedrijven Financiële Luchtfietserij,' *NRC Handelsblad*, 13 januari, Rotterdam.

Wallis, K.F., 1993, 'On Macroeconomic Policy and Macroeconometric Models,' *The Economic Record*, 69, pp. 113-130.

WRR, 1987, *Ruimte voor Groei; Kansen en Bedreigingen voor de Nederlandse Economie in de Komende Tien Jaar*, Rapporten aan de Regering nr. 29, The Hague, SDU Publishers.