<u>On the usefulness of macroeconomic models</u>

by

Professor Angus Deaton
(University of Bristol)

Contents

On the usefulness of macroeconomic models

Introduction

1 I should like to begin with a disclaimer. I am not a macroeconomist, however that may be defined, nor are my econometric interests primarily directed towards the construction and operation of macroeconometric models. However, I have had a good deal of contact with such models, through my association with the Cambridge growth project, through my (past) membership of the Treasury Academic Panel, and through occasional contact associated with my own work (which has from time to time touched on areas of interest to macroeconomic modellers). Even so, compared with the other contributors to this discussion, I am little more than an interested bystander and my knowledge of all the relevant literature is so sketchy as to inevitably lead me into error. However, involvement in macroeconometric modelling seems to be an all-consuming activity, so that perhaps a sympathetic outsider is well-placed to evaluate both the position of the modellers and the dissatisfaction sometimes expressed by sections of both the public and the profession.

2 It is a truism that if macroeconomic models could do half of everything that has ever been claimed for them, their payoff would be so large as to justify the commitment of resources well beyond anything they are ever likely to obtain. Yet a potential payoff must be assessed in relation to the probability of success. In my notes below, I try to summarize what I believe that macromodels can and cannot do, as well as the fundamental technical reasons which limit their scope. The first section deals with some of the econometric issues of model building and model testing and if the picture seems somewhat bleak, it is so only in relation to what I feel are hopelessly optimistic expectations. The second section is concerned with what I call the 'social' role of model building and attempts to assess the importance of the models within the policy-making and academic establishments. I attempt no overall conclusions.

Econometric aspects of macroeconomic models

(a) Model structure

3 The ideal framework for macroeconometric modelling would be a universally accepted theoretical framework within which the role of

the econometrician would be simply to assign (or infer) values for the parameters. Progress would consist of inventing more efficient estimation techniques and obtaining more data, both of which would be reflected in increasing precision of estimation as time passed. This sort of situation is well described in most econometric textbooks. These agreed models would be suitable objects for optimal control, and economic policy making would be a straightforward matter of social engineering. Why is it that this picture is so far removed from what actually takes place? Clearly, the structure is not known, and it is not only a matter of settling details. As recent debates have made clear, the theoretical foundations of macroeconomics are subject to more dispute now than has ever been the case since econometric model building got seriously under way. Perhaps the deepest division is between those who view modern economics as essentially equilibrium structures and those who do not, but there is room for a large number of variants and mixtures, particularly since the disequilibrium models are very far from being coherently worked out, even in theory. Policy-makers (not to mention the public) are understandably impatient with theoretical debate of this sort; after all, scientists can confer all the benefits of technology without agreeing on the fundamentals of general relativity or the basic composition of matter. Why cannot econometricians do the same? Surely econometric models, together with the tools of modern statistical inference, can test out the various theories, discard those which contradict the data and settle on a viable framework for the conduct of economic policy? I shall argue below that such goals are much too optimistic.

4 There are, of course, those among us who are not troubled by uncertainty as to fundamental model structure. And indeed, at some level, all econometric models must be based on an untested (and within the model, untestable) system of postulates. For some model builders these postulates are 'obviously' correct; to some the world is 'clearly' in disequilibrium, to others it is 'obvious' that relative prices are not very important, and there are those who regard with horror any suggested deviation of reality from a perfect Walrasian system. The very diversity of these untested model constructs demands a high degree of scepticism from anyone who, like me, is unfortunate enough not to have been shown truth by divine

revelation. However, it is not true that models embodying controversial assumptions are useless and there is certainly a case for allowing individual modellers to follow up their theories. However, it is of the greatest importance that the results which come out of these models are correctly interpreted. A distinction must always be drawn between those results which are genuine 'findings', are data-dependent, and which could conceivably have come out differently, and those others which are not results at all but are the inevitable consequences of the assumed model structure and which are independent of any possible values of the parameters. A model which allows little influence for relative prices cannot 'find' that the control of the economy requires the imposition of quantity constraints; it assumes it. Nor can an equilibrium model of the British economy 'show' that unemployment is voluntarily chosen leisure. Although the distinction between findings and assumptions is an elementary one, it is rarely emphasized in reporting macroeconometric results. No doubt this is in part because the models are so complex that the distinction can be difficult to make, but it is hard to escape the feeling that model operators are sometimes unaware of which is which. It is very easy when working with a model on a daily basis to forget that the model and reality are two different things. And to the extent that the audience is imperfectly informed, the output of a large 'scientific' econometric model is more credible and impressive than the crude assumptions which are its input.

(b) Model testing and comparison

5 Why then is it optimistic to expect econometric models to compete amongst themselves leaving only the fittest to survive? And why, even when the models are not different in fundamentals (e.g. conventional IS-LM Keynesian models), and when the parameters are estimated using much the same techniques on much the same data, do the results and predictions differ so much?

6 The most popular criterion for evaluating model performance is the forecasting record and some would claim this is the only legitimate criterion. As is well known to anyone who has ever tried to conduct comparative tests of track records, there are considerable practical difficulties in implementation. Data are revised so that past

predictions are of magnitudes which officially no longer exist. Similarly, exogenous variables take unexpected turns and it is hardly legitimate to dismiss models for failing to predict variables which they make no claims to be able to predict. Even so, there are more or less satisfactory solutions to these problems and I shall ignore them from now on. The more fundamental problem with the forecasting criterion is that there is no reason to expect nature (or indeed the government) to conduct the crucial experiments which are necessary to distinguish between different models, or even minor variants of the same model. If the future has much the same general structure as the past, false models will predict as well over that future as they did on the past. Success in forecasting tests therefore means very little. Failure, of course, means just that, although experience suggests that at times of major structural change, predictive failure is not confined to one or even a subset of models. One is then driven to the clearly unsatisfactory procedure of selecting the model which fails least dramatically. Such experience suggests that 'laboratory' experiments, not real ones, are required and that models ought to be subjected to severe diagnostic testing during construction in the hope of detecting and correcting what will otherwise later appear as predictive failure. Hopefully such diagnosis can use more powerful and more sharply discriminatory tests than will be delivered by unassisted nature.

7 How then should diagnostic testing proceed and is it possible to ascribe the supposedly unsatisfactory performance of models to insufficient or incompetent testing? I think that it is part of the story but it is not the major problem. It is clear that standards of practice in this respect have improved enormously and are still in the process of improving. In the days of the first econometric models, formulations had to be both parsimonious and simple, and if theory was to be useful, it had to be taken for granted and not tested. Today, data is relatively more plentiful (although still pitifully short by the usual statistical standards) and formulations can be more flexible and more ambitious. In particular, the simpler structural relationships can be tested against more general alternatives, and those which survive a battery of such tests are worthy of a good deal of confidence. Much recent work in econometrics has gone into the design of diagnostic tests for various circumstances and much has

been learnt about appropriate techniques of econometric analysis on time series. There is also no doubt that the use of these tests and procedures has improved and will improve further the quality of (at least some of) the structural equations of macroeconomic models. Or, to put it at its most negative, current equations are less often implausible, data-inconsistent or unstable than they were a few years ago. This is undoubtedly progress and is much to be encouraged.

8 Yet the application of improved testing of structural equations has limitations of its own as a methodology, and, in my view, it does not address the crucial issue for policy formation. I take the methodological limitations first. These are largely, but not exclusively, the result of working with very small data sets, with rarely more than 80-100 observations. First, the asymptotic theory on which many of the diagnostic tests are based may not be very useful in such samples, so that correct models may be too frequently rejected, false ones rarely detected, or both. Second, testing usually consists of comparing the equation being considered with some more general alternative. But with short time series, general alternatives are estimated with little precision and are not robust with respect to errors in variables, data revisions, changes of sample period, and so on. The statistical rejection of a theoretically plausible specification against a largely meaningless alternative is thus subject to a number of different interpretations. Which leads into the third difficulty. Statistical inference is at its best when asked to discriminate between well-defined theoretically coherent alternatives. Non-nested tests can be used, and in the standard nested case, the generalised alternative is itself a proper theoretical construct. But for many structural equations in a macroeconomic model, the theory is weak. Obvious examples are some of the equations required to close the system (e.g. the relations which transform government expenditure, company profits and wage bills into disposable income), although there are many well-studied relationships where theory is deficient. Faced with this, testing is frequently against atheoretical specifications, the serious adoption of which would make a nonsense of the whole model structure, not to mention its overall properties. Data-based theory construction is dubious and difficult at the best of times; with short time series it is just not viable, except possibly for very simple relationships. The consumption

function is a good illustration of the point; recent work has suggested that a data-based approach can yield useful results. Yet, even here, the fifty years of theory behind the relationship plays a crucial (if implicit) part, and the inferential procedures are greatly eased by the essentially bivariate nature of the relationship.

9 For policy use of econometric models, structural equations are merely building blocks and are of little interest in their own right. What is required for policy analysis is the relationship between the exogenous variables, including policy instruments, and the endogenous variables of the system, particularly those which are of direct policy concern. Even for forecasting, it is these reduced form relationships which count, in the sense that we want to know that feeding-in actual exogenous variables will generate a useful approximation to actual endogenous variables. Since we would automatically distrust forecasts which turned out to be correct because of large offsetting errors, it is hard to see any operational difference between forecasting and policy analysis requirements. In either case, the usefulness of the model is directly related to the accuracy, robustness and stability of its reduced form (or dynamic equivalent) and not its structural equations. Of course, if the structure were perfect, so too would be the reduced form. But we know that all structures are at best more or less accurate approximations, and it is quite possible (indeed it is likely) that tolerable errors in the structural equations will induce intolerable errors in the reduced form.

10 If we are to fully understand the properties of a macroeconomic model, and if we are to diagnose its possible faults and deficiencies as a policy tool, it is necessary to understand and be able to test its reduced form. And this is where the principal difficulty lies. Consider first two cases where the matter is straightforward - small analytically soluble models, and the classical linear simultaneous model. In the former, the solution is derived algebraically and its properties can be studied directly; these will often depend on the precise configuration of structural parameters (e.g. for stability), but given parameter estimates (or even rough ranges) the complete behaviour of the model solution is essentially known. The construction

and study of such models is the essential subject matter of all macroeconomic theorizing, and virtually all the insights we possess about the behaviour of macroeconometric systems come from such models. Their amplification into large scale macroeconomic models is clearly necessary both for detail and for the realism which the small (largely academic) models lack. But I can think of almost no instances where the large model illumines something that is not already clear in the small 'canonical' model which lies at its heart. The other straightforward case is the text-book model usually written as:

$$By + \Gamma x = u \tag{1}$$

where y is a vector of endogenous variables, x is a vector of exogenous variables, u is a vector of disturbances, and B and Γ are parameter matrices. (Obviously realism would require the insertion of dynamics, but the simple form, provided it remains linear, will illustrate my point.) Equations such as (1), like the small models, are directly soluble to give:

$$y = -B^{-1}\Gamma x + B^{-1}u. \tag{2}$$

Note that size is no difficulty for solving, only nonlinearities. Equation 2 illustrates immediately the point that small errors in structure may have far from small consequences in the reduced form; matrix (like scalar) inversion transforms small magnitudes into large ones. But, given linearity, untoward consequences can be avoided by estimating equation 2 directly, for example by maximum likelihood techniques. This guarantees that the responses of y and x are not absurd, at least over the sample period, and such methods have the dual advantage of estimating directly the policy-relevant parameters ($\partial y/\partial x$) while incorporating the structural information that these should take the special form $-B^{-1}\Gamma$. Diagnostic checks can easily be incorporated, most obviously by testing whether the unrestricted regression of y on x yields coefficients not grossly in violation of the structural restrictions. Suspect response coefficients are easily identified and checked so that, given a thorough (but essentially routine) econometric review, the responses from such models can be accorded some confidence. Contrast this with the reduced form responses obtained from a model constructed by paying attention to only the structural equations where the consequences of either over-zealous theory or over-zealous econometric testing may be parameter estimates which in the reduced form have nonsensical consequences for prediction and policy making. Putting the same point in more Bayesian terminology, the parameters of interest are $(-B^{-1}\Gamma)$, not B and Γ separately, so that estimation with respect to the former more accurately reflects the loss function of the model user. (I note parenthetically one final point. Fully efficient estimation of equation 2 requires knowledge or estimation of the variance-covariance matrix of $B^{-1}u$. For large models this is typically impossible on normal length time series. One is then forced to make essentially arbitrary assumptions about the error structure to achieve identification of the parameters. Asymptotically, this makes no difference, but we know very little about the effects in the sort of samples we actually deal with.)

11 The standard macroeconometric model of today is not small, it is not analytically soluble, and it is certainly not linear. Beyond the essentially general structural form:

$$f_i(y, x, u) = 0, \qquad (3)$$

very little can be said. How can the properties of such models be understood and how can we check whether or not they make sense? Note that the reduced form corresponding to the structure equation 3 cannot normally be written down in closed form, i.e. with y as a function of x and u. Nor indeed do the policy-essential derivations $\partial y/\partial x$ exist in the usual sense. If equation 3 is correct, and even if all the u's are 'nicely' behaved, the responses are random variables which typically do not even possess moments; it thus makes no sense to talk about the average or expected response of an endogenous variable to a policy instrument. Certainly, having estimated the model, it is possible to solve out values of y for given x computationally on the assumption that u is zero, and this is frequently done. What is unclear is what such values and the responses implicit in them are supposed to mean; they certainly cannot be accepted as typical of the actual behaviour of the model. My belief is that the complete reduced form properties of large, non-linear models are both unkown and unknowable and that, *a fortiori*, it is impossible to subject them to the sort of diagnostic testing which would prove whether or not they are suitable vehicles for policy analysis. Model proprietors will challenge this; they, if no one else, understand the behavioural characteristics of their creations. But even if this knowledge is real (and for the reasons given I doubt it), information which is revealed only to high priests has always provided a somewhat dubious oracle for the rest of mankind. I have on many occasions been 'impressed' by the ability of model builders to tell plausible economic stories about the most apparently implausible responses, including even those later traced to computational or typographic errors. The properties of econometric models can be fully understood if, and only if, they are analytically soluble. Hence, if macroeconometric models occasionally produce predictions or policy responses which are in prospect implausible and in retrospect silly, it is not necessarily due to econometric incompetence or incomplete checking; an appropriate methodology just does not exist.

(c) Stability and instability

12 It is perhaps worth relating the foregoing arguments to the currently much-debated stability critique of conventional macroeconometric models. In the final analysis, it seems to me that stability or not is an empirical matter; rational expectations or not, information is far from perfect and some behavioural rules are costly to adjust. In any case, the stability or instability of structural equations can be tested straightforwardly by the techniques discussed in paragraph 8 above. While it is true that many of the macro models existing in the late sixties later showed evidence of instability, very few of these had been adequately tested over the sample period for parameter instability, so that their subsequent failure is just as likely to have been due to (by current standards) poor econometric technique as to some fundamental and inevitable instability resulting from a change in perceived policy rules. Which actually occurred should be amenable to empirical testing, and my impression is that instability of the underlying relation is less of a problem than inadequate or incorrect econometric representation, although further findings may change this view. However, if the stability critique does turn out to be empirically important, the question arises as to how models should be built which are likely to be robust. An increased emphasis on expectations is clearly a good

idea from many points of view and should not be controversial. However, there is also the argument that models should be much more explicitly based on the specification of preferences and technology since these are the real foundations on which economic theory is built and only they can be expected to be stable. Whatever view one takes on the appropriate basis of economic theory, this argument seems mistaken even on its own terms. Empirically, preferences and technology are represented by (fairly general but always) local approximations so that, except in the highly unlikely event of the approximation being exact, the parameters of the approximation will change as the economy moves, just as do all the other parameters of the system. Hence, if the stability critique is really valid and of quantitative importance, we can give up altogether; otherwise we should adopt the eminently sensible alternative of testing carefully for instability.

(d) Some recommendations

13 Before going on to argue the case for macroeconometric models in the next section, I wish to summarise the positive recommendations from this section, and also to make one new point. Summarising first:

(i) I believe that small, analytically tractable models are still relatively underused as a guide to policy and that it is easy to overestimate the gain from complexity and size. On the other side, small, explicit models are a great advance over the small, implicit and possibly self-contradictory models which tend to be carried around in people's heads.

(ii) Although limited in scope, diagnostic tests of structure ought to be pursued as far as they can be given the data available.

(iii) Expecting too much from large models is not productive and there are good technical reasons for them not achieving what may be their most obvious and explicit aims.

14 Finally, I should like to argue strongly for a rather different emphasis in applied macroeconomic research. Nearly all of the problems discussed in the previous paragraphs would be greatly eased

if we had a better understanding of how the economy actually worked, of how firms and consumers actually respond to policy and other changes. Aggregate time series data are a very poor tool for the testing and building of proper theory. I believe that very great progress could be made, in some directions at least, if proper panel data were available. As it is, there are virtually none in the United Kingdom on either households or firms (at least that are available to researchers). The United States has recently started a rotating quarterly panel of households and the possession of such data here would enable real scientific progress to be made on topics of fundamental interest to macroeconomic modelling. Inevitably, the payoffs could not be immediate, but without this type of research, macroeconomics is going to remain as it is now, at the mercy of academic and political fashion.

Social aspects of macroeconomic models

15 In the previous section, I argued that little in the way of scientific knowledge is to be gained from the construction of large-scale models over what can be learnt by other means. At present at least, there are very few spin-offs into academic advance, nor in either this country or in the United States are the very best young economists being attracted to work on the model-building teams. However, fashions are susceptible to change and it is certainly true that excellent research training has been provided within the teams in the past. Even so, I think the positive arguments for the existence of the models are to be found elsewhere.

16 First, there is no doubt that their existence has raised the standard of the policy debate in this country. Even if incorrect, they provide a coherent, necessarily consistent, and essentially public framework within which debate can take place. Without this, policy is entrusted to 'wise' men who are not accountable to external public or academic criticism. In the atmosphere provided by the models, policy formation is more answerable to informed opinion. In practice, of course, the changes have been marginal but sure 'both significant and of the right sign'.

17 Second, it is almost certainly desirable to allow people who wish to elaborate new systems of representing the economy to do so. As

argued above, the elaboration and enlargement teaches very little beyond the insights given by the small models on which new approaches are inevitably based. But a full-scale representation is necessary if all the detail is to be filled in, and if outsiders are to have the opportunity of informed criticism. The blueprint may turn out to be seriously inadequate when built at full-scale and the attempt itself generates activity and stimulates debate. Even so, experience suggests that outside pressure is continuously needed if model builders are to seriously compete and debate with one another. This may come naturally from the flow of economic events, but there is always a tendency for each team to cultivate its own garden as if the others did not exist. In particular, it seems possible for models and model-teams to exist for long periods of time without being forced to meet even valid and well-thought out academic challenges. Political protection has often seemed to be more important than the ability to meet peer-group review.

18 In all, I find that an overall conclusion does not readily suggest itself. I think a more realistic and less ambitious view of modelling ought to prevail both among policy makers and among the model-builders themselves. But that is far from saying that the models are of little use and ought not to exist. Indeed, it is very hard to see any other way of conducting an informed debate about the realities of economic policy.