The development of quantitative empirical analysis in macroeconomics

by

Fernando Ballabriga
The development of quantitative empirical analysis in macroeconomics*

* The author acknowledges comments from André Dramais and Lars Jonung, DG ECFIN
Abstract

This paper contains a schematic description of selected relevant events in the development of quantitative macroeconomics since the start of the Cowles Commission for Economic Research in 1932. It also provides a sketch of what could be a promising path for future events.
1. Introduction

Is it possible to find empirical counterparts of demand and cost functions? Can we obtain quantitative laws of motion which characterise the evolution of macroeconomic aggregates, such as the levels of production and employment? Or to go even further, can careful empirical research provide quantitative guidelines for economic policymakers whose decisions aim to influence the evolution of their countries’ economies?

Finding answers to this type of questions should be the ultimate, perhaps the ideal, goal of a science like economics which has a strong quantitative component and aims to contribute to the improvement of social welfare.

However, up until well into the 20th century, Economics was essentially a qualitative field of study. There was naturally an awareness of this shortcoming, which led economists such as Irving Fisher, Ragnar Frisch, Joseph Schumpeter and others to found the Econometric Society in 1930 in order to promote the development of the quantitative side of economic theory. Three years later, in 1933, the society published the first issue of its journal *Econometrica*, which featured an editorial that advocated bringing together economic theory, mathematics and statistics as a key strategy for progressing in the development of quantitative economics.

How far has this development come? What is the outlook for further advance in the near future? The following sections elaborate on these questions focusing on a selection of relevant events in the development of quantitative macroeconomics.

2. Conventional Econometrics

Shortly after the Econometric Society was founded, Alfred Cowles contacted some of the original members. Cowles was president of an investment
consultancy and was interested in forecasts of stock market evolution and, by extension, in economic research. The contacts proved fruitful. 1932 saw the start of the Cowles Commission for Economic Research, which was made up of members of the Econometric Society (with which it shared its headquarters until 1955) and financed mainly by Cowles.

The Cowles Commission was based in Chicago from 1939 to 1955. During this time, and particularly during the 1940s, Commission members laid the groundwork for what could be referred to as “conventional econometrics”, making two basic contributions to the field: they advocated the use of statistical inference in economics and developed simultaneous equation models up to an operational stage, dealing with their identification, estimation and validation.

There are inherent links between these two contributions. On the one hand, using statistical inference made it necessary to visualise economic data as generated by a random process. On the other hand, specifications for this random process had to involve simultaneity because it is a key feature of economic interactions. The challenge therefore lay in how to use appropriate statistical techniques to obtain a quantitative version of random simultaneous equation systems which would reflect the way the economy actually works.

Simultaneous equation systems include up to three types of structural equations: identities, technological restrictions and rules governing the behaviour of economic agents. The existence of simultaneity causes problems of observational equivalence and estimation which do not usually occur when simultaneity does not exist. The Cowles Commission addressed the problem of identification by resorting to economic theory and the use of exclusion restrictions when specifying econometric models. In addition, it set the basis for developing the least squares and the maximum likelihood estimation methods which commonly appear in present text books.

For three decades the Cowles Commission’s econometric principles were the framework of consensus for the profession and monopolised econometric
theory and practice. In the specific field of macroeconomics, Klein (1947) was the first to construct macroecoanometric models which could be used in economic policy decision-making processes. The size of these models gradually increased. They began to be systematically used to quantify the macroeconomic impact of different scenarios defined in terms of alternative paths for the model’s exogeneous variables. During recent years particular efforts have been made to include rational expectations and arrive at a detailed modelling of international connections. Nowadays one of the most frequently cited models is the Wharton Model, of which there are short and medium-term versions which include more than one thousand equations.

3. Critiques of Conventional Macroeconometrics: Lucas and Sims

In the second half of the 1970s, two authors wrote articles criticising the uses and basic principles of conventional macroeconometric models: Lucas (1976) and Sims (1980). Both critiques of the conventional modelling strategy were so thorough that, according to the authors, it would be advisable to discard this strategy and seek alternatives that would correct what they considered to be unacceptable features of the conventional methodology. Indeed, their articles were very influential in the United States, triggering the start of the research programmes suggested by the authors.

Lucas and the econometric programme of the rational expectations’ school

Lucas’s critique was perhaps the most revolutionary one. Using arguments based on the rational expectations hypothesis he dismissed the use of conventional models in exercises involving assessment of alternative economic policies.

More specifically, consider the following econometric model:

\[\text{\textsuperscript{1}} \text{Taylor (1993) is a good example of this line of work.}\]
\[ Y_1 (t) = F (Y_1 (t - s), Y_2 (t - s), s \geq 0 ; \delta_P) + u_1 (t) \]  
\[ Y_2 (t) = G (Y_2 (t - s), Y_1 (t - s), s \geq 0 ; \delta_G) + u_2 (t) \]

where \( t \) is a time index, \( u_1 \) and \( u_2 \) are model disturbances, \( \delta_F \) and \( \delta_G \) are parameter vectors and, for the sake of convenience, the model's variables have been separated into a \( Y_1 \) vector that represents the private sector, and a \( Y_2 \) vector for economic policymakers. Given \( G \), Lucas argues that expectations make \( \delta_F \) a function of the parameter \( \delta_G \) and of the \( \sigma \) parameters that characterise the distribution of the disturbance vector \( u_2 (t) \):

\[ \delta_F = f (\delta_G , \sigma , \delta_P) \]  

In other words, the specification of private sector behavioural equations depends on the control variable generation process (as well as on the \( \delta_P \) parameter vector which we will discuss shortly). Thus, to the extent that the alternative scenarios considered for \( Y_2 \) involve changes in \( \delta_G \) and/or \( \sigma \), evaluating economic policies in the conventional way, i.e. projecting with a fixed \( \delta_F \), to determine the behaviour of the private sector as regards these alternative scenarios is incorrect inasmuch as the estimated equations for the private sector are no longer valid.

A different strategy is required for relevant analyses. This involves estimating the “deep” parameters. In other words, estimating the \( \delta_G \), \( \sigma \) and \( \delta_P \) vectors of parameters, the latter representing agents' preferences and technology. The vector \( \delta_P \) is assumed to be independent of the stochastic laws that generate the control variables (\( \delta_G \) and \( \sigma \)), and appears in (2) by virtue of the optimising behaviour of the economic agents. Once these parameters have been estimated, it is possible to make legitimate policy evaluations by combining (1)
and (2) because the effect of a change in scenario on the process that describes private sector behaviour is taken into account. This is the underlying philosophy of the school of rational expectations’ econometric programme, whose final objective is (Sargent, 1984) to seek policy actions which generate the most desirable stochastic process for the economy. The aim of this search is to be able to offer quantitative advice in order to shape government actions for the years subsequent to the sample period.

The strategy used to obtain the “deep” parameters involves specifying and solving dynamic optimisation problems under the hypothesis that agents’ expectations are rational. This framework of analysis is adopted with the aim of interpreting the correlations that usually characterise time series of aggregate economic data whilst respecting the basic theoretical principle that the agents’ observed behaviour changes when there are changes in the restrictions they face (this is the essence of Lucas’s critique).

Solving the optimisation problem generates stochastic laws of motion for the variables studied which depend on the parameters that characterise the economic structure used. These “deep” parameters are then estimated and the underlying economic structure tested. An important consequence of adopting this framework of analysis is that identification restrictions are available in the form of non-linear functional relations between the coefficients of the optimal stochastic laws of motion, which, to some extent, reduces the need for exclusion restrictions which are typical of conventional econometrics.

Sargent (1981) and Hansen and Singleton (1982, 1983) are classic references in the econometric programme of the rational expectations school which emerged subsequent to Lucas’ critique.

**Sims and the VAR Methodology**
C.A. Sims’ macroeconometric proposal stemmed from a direct criticism of the methods used to construct conventional models. A description of his arguments follows.

The validity of the restrictions used to obtain a structural interpretation is crucial if you aim to defend the idea that there is some connection between reality and the model used to represent it. Sims argued that the majority of restrictions used to identify conventional macroeconometric models are unbelievable. They are not justified by economic theory. Indeed, theory does not provide sufficient unequivocal restrictions relative to the number of variables and equations usually included in conventional models. In particular, the exogeneity of many variables is fictitious rather than real.

Let us look again at the econometric model (1) for purposes of illustration. If, as is usually assumed in economic practice, F and G are linear, this model suffers from an identification problem because the two equations are statistically indistinguishable, making it impossible to decide which of them reflects private sector behaviour and which reflects the behaviour of economic policymakers. In order to solve this problem, it has been common practice in conventional modelling to treat the control vector as exogeneous. In other words, to reduce equation (1) to the following restricted specification:

\[ Y_1(t) = F(Y_1(t-s), Y_2(t-s), s \geq 0; \delta^*_F) + u_1(t) \]
\[ Y_2(t) = G(Y_2(t-s), s \geq 0; \delta^*_G) + u_2(t) \]

where the \( Y_1 \) vector has been eliminated from equation G and the vectors of disturbances \( u_1(t) \) and \( u_2(t) \) are assumed to be orthogonal. The exogeneity of \( Y_2 \) clearly guarantees the identification of the F and G equation blocks, but it is very likely that this is an unjustified assumption inasmuch as policymakers responsible for controlling \( Y_2 \) usually react to the private sector events reflected in the evolution of \( Y_1 \).
Sims maintains that when model identification rests on such a fragile base its implications in terms of the economy’s underlying interrelations can scarcely be taken into consideration, which disqualifies it as a tool for empirical analysis.

The methodology proposed by Sims (1980) involves specifying and estimating macroeconometric models that do not include a priori controversial restrictions. In fact, he proposed specifying minimally restricted models in which all the variables with a clear economic content would be treated as endogeneous. The resulting models are known as Vector Autoregressions (VAR). Models of this kind are obtained from (1), assuming that $F$ and $G$ are linear and solving for the contemporaneous value of the endogeneous variables:

$$Y_1(t) = F(Y_1(t-s), Y_2(t-s), s > 0; \beta_F) + \varepsilon_1(t)$$

$$Y_2(t) = G(Y_2(t-s), Y_1(t-s), s > 0; \beta_G) + \varepsilon_2(t)$$

(4)

Under the assumption that the vector of stochastic disturbances $(\varepsilon_1, \varepsilon_2)$ is white noise, (4) would be the VAR representation of the vector of endogenous variables $(Y_1, Y_2)$.

The implementation of Sim’s proposal soon found obstacles which ended up becoming subjects of discussion and research throughout the 1980s and 1990s. The first obstacle was the wide parametrisation of VAR models. The second was the absence of a specific identification proposal, making VAR models reduced-form models with no economic interpretation. Nowadays, both the problem of degrees of freedom and the problem of identification have been solved in a relatively satisfactory way, which has facilitated the spread of VAR methods and the understanding of their underlying motivation: the acknowledgement of the fact that there is widespread uncertainty about the economic data generating process.

The immediate consequence of this acknowledgement is that an appropriate modelling strategy should explicitly include this uncertainty in the model.
specification process in order to treat it systematically and objectively. It is precisely this idea that justifies Sim’s insistence on keeping restrictions to a bare minimum so that relevant empirical regularities can be extracted by giving economic data the most objective reading possible. Development and instrumentation of this modelling strategy was accompanied by the introduction of Bayesian statistic techniques which have become one of the distinguishing features of the VAR methodology.

In addition to Sims (1980), other classic references on VAR methodology are Doan, Litterman and Sims (1984), Sims (1986), Bernanke (1986) and Blanchard and Quah (1989).

4. The Econometrics of General Equilibrium

Although based on severe criticism of conventional econometrics, Lucas’ and Sims’ proposals both share its basic premise that specification, estimation and statistical testing of equation systems is a valid procedure for comparing economic theories. However, Kydland and Prescott (1991a, 1996) have recently questioned their validity.

Kydland and Prescott adopt Lucas’ (1980) proposal that an economic theory is an explicit set of instructions which generates time series for economic variables; a model with economic agents which correspond to their real-life counterparts (consumers, businesses, government); a computerised picture of the national economy. Their argument is that, according to this definition, the equation systems used in modern econometrics are not economic theories but only simple sets of statements about how the economy works. Inasmuch as testing a theory means that it must first be formulated, the authors conclude that the statistical test of hypotheses which is customary in econometric practice is not an appropriate tool for testing economic theories.
Their alternative proposal for testing economic models is what they call the computational experiment, a term borrowed from physics in which this validation method is commonly used.

Generally speaking, the computational experiment involves obtaining the stochastic laws of motion for the variables under study by using the theory you want to test and then using a computer simulation of these laws to obtain the quantitative implications of this theory. In more concrete terms, and as regards the specific field of economics, Kydland and Prescott describe the computational experiment as an exercise that takes the following basic sequence:

1. Ask the quantitative question you want answered.

2. Construct an economic model using well-tested theory, i.e. theory that has been previously tested and provided satisfactory answers.

3. Assign numerical values to the model parameters so that their simulations reproduce clearly established empirical regularities. This assignment is called “model calibration”.

4. Perform the experiment with the calibrated model. In other words, simulate the model in order to obtain an answer to your quantitative question. The more sensitive your simulation is to the model calibration, the less accurate and answer will be, and vice-versa.

A recent field of study whose development is based on computational experiments is the Real Business Cycle theory proposed by Kydland and Prescott (1982,1991b). The specific question they pose is how much the post-war US economy would have fluctuated had technological shocks been the only source of macroeconomic variability. They used the neo-classical growth theory to construct the model. The model is calibrated with the aim of reproducing some of the average ratios observed in the post-war US economy, among
them, the ratios of consumption and investment to GDP, the ratio of capital income and labour income, and the ratios of number of hours worked per employee and number of employees to the total change registered in the number of hours worked. The outcome of their experiment revealed that the variance of the simulated output is 70% of that of the observed output. This estimate proved robust when various changes were made in the original model.

Inasmuch as statistical inference is not inherent to the computational experiment, the question that immediately arises is whether the computational experiment is really an econometric tool. Its advocates insist that it is. In fact, Kydland and Prescott claim that the original meaning of the term “econometrics” coined by Frisch and clearly explained in the editorial introducing the first issue of *Econometrica*, is to derive quantitative implications from economic theory, which is precisely the objective of the computational experiment.

The authors continue by pointing out that the meaning of the term contrasts with current econometrics meaning and practice, which involves estimating economic relations based on goodness of fit statistical criteria. Kydland and Prescott stress the distinction between “estimation” and “calibration” as the differentiating feature between the econometrics of the computational experiment and the more popular version of econometrics which they call “systems of equation econometrics”. This approach to econometrics determines the value of equation parameters in order to obtain the best fit to the variables (estimation). In contrast, the computational experiment technique selects the theoretical model constraining its parametrisation with the aim of reproducing certain well-established stylised facts (calibration).

In the authors’ opinion, the goodness of fit strategy is incorrect because it is the same as evaluating the explanatory power of certain variables over others on the grounds of a parametrisation that is deliberately selected in order to maximise this very explanatory power. They maintain that the model should be selected in accordance with the quantitative question to be answered and existing economic theory, and that the credibility of the answer obtained with the
selected model should not depend on statistical test and fit, but on the validity of the underlying theory. This validity should be based on the model’s capacity to reproduce well-established empirical regularities. This ability is suitable for testing models.

5. The Sceptical View: Scientific Illusion in Empirical Macroeconomics

When economics and natural sciences are compared, one is struck by the dichotomy between theoretical and empirical economics. This more or less explicitly corresponds to a classification whereby theoretical economics is considered a first division player and empirical economics is a second division player. This classification tends to be based on the perception that formal econometric studies have had practically no impact on the development of economic science and leads to a sceptical question of whether the large amount of resources that has gone into developing formal econometrics has even been worth the effort. Rather than bringing together a specific line of thought, this vision can be better described as the shadow of a doubt that hovers over the profession. It was approached from a particularly nihilistic angle by Summers (1991), whose arguments are summed up below. He compares formal econometric analysis with its pragmatic and informal counterpart and opts for the latter.

In his description, Summers identifies formal econometrics with analyses that resort to statistical inference in order to estimate structural parameters, test hypotheses derived from economic theory, and isolate causal relations in systems with numerous interdependent variables. He points out that, contrary to the case of empirical work in physics, formal econometric studies contribute little or nothing to scientific progress in economics. According to the author, this is revealed in a number of different ways: important theoretical progress in the past few decades has not required formal econometrics, there is no interest in replicating their findings even when this is possible, nor can one recall any formal econometric study which has made a substantial contribution to
economic science. Summers believes that formal econometrics has so little impact because it mixes up methodological and substantive contributions, priding itself on sophisticated methodology when significant empirical research should be based on new information, not new techniques.

In order to illustrate his point, Summers examines empirical work which has been classed as relevant in each of the two tendencies he considers dominant in formal econometrics: the rational expectations school and VAR macroeconometrics.

The work he selected from the rational expectations school’s econometrics programme is that of Hansen and Singleton (1982) which attempts to shed light on how the price of assets is determined. For that purpose these authors specify a representative agent macroeconomic model, estimate its deep parameters, and ask whether the model is or is not statistically rejected by the data, concluding that it is.

Summers criticises this work for failing to contribute anything at all. On the one hand, simply concluding that the data reject a model which you already knew to be false (as are all models) is not significant unless the exercise at least explains the reasons for the deviation of data from theory, thereby suggesting further streams of research, as, for example, Mehra and Prescott (1985) did when they suggested that rationalisation of the high risk premium observed probably requires the Arrow-Debreu framework to be abandoned in favour of an incomplete market framework.

On the other hand, due to the tremendous uncertainty that characterises the model specification and, by extension, its parameters it is doubtful that anyone would be inclined to use Hansen and Singleton’s deep parameters to forecast the effects of a possible intervention in the economy. Summers concludes that Hansen and Singleton’s structural estimation and statistical test are no more than examples of methodological elegance.
His conclusion is equally negative when he analyses Bernanke’s (1986) work in the field of VAR macroeconometrics, seeking an empirical discrimination between the alternative explanations of the positive correlation observed between aggregate production and money. In his opinion, any one of the directions of causality established by Bernanke can be interpreted in terms of inverted cause-effect or plausible alternative, which means that his results do not establish empirical regularities that could help clarify the traditional debate about the nature of the correlation between monetary and production aggregates. For example, Bernanke concludes that credit is a relevant factor in explaining cyclical fluctuations, but nothing in his model denies the possibility of an increase in credit in anticipation of an expected increase in economic activity caused by other factors. Generally speaking, Summers considers it fruitless to attempt to identify directions of causality among variables by resorting to sophisticated methods of statistical analysis and without introducing information other than that contained in repeatedly analysed sets of time series.

Summers compares the limited influence of formal econometrics with the influence of natural experiments along the lines of Friedman and Schwartz (1963) or, in more general terms, the informal pragmatic empirical analyses contained in Modigliani and Brumberg (1955), Phillips (1958), Feldstein (1974), and others. Though this is a simple analysis, it is capable of establishing clear empirical regularities that stimulate theoretical analysis and lead to progress in understanding how the economy really works.

Here one is struck by the overwhelming influence of Friedman and Schwartz’s monetary history on discussions about the macroeconomic effects of money; the conceptual advance and subsequent research stimulus provided by Modigliani and Brumberg’s evidence that wealth is an important factor in explaining consumer spending; the continued timeliness of the Phillips curve; or the discussion triggered by Feldstein’s evidence of the impact of social security on private saving.
In Summers’ opinion an empirical analysis that is able to establish stylised facts that stimulate explanation is essential if theorists are to be convinced to take the evidence seriously, reduce their excessive emphasis on internal consistency, and put an end to the present lack of connection between the two fields of work.

6. Current Situation and Future Prospects

The foregoing description clearly reveals that the field of empirical macroeconomics is currently in disarray. There are a number of different tendencies which are difficult to reconcile. Nevertheless, if we look closely at the above we can manage to identify several contrasting positions which, if brought together, could mark the start of a path towards future consensus. Specifically, there are three important and closely related contrasting positions that underlie the foregoing discussion, and I will conclude this paper by discussing them: (a) the contrast between natural and social sciences, (b) between descriptive statistics and statistical inference and (c) between Classical and Bayesian inference.

Economics is probably the social science which has relied the most on the learning methods used in the natural sciences. However, as a social science, it does not have recourse to controlled experimentation. Therefore the mechanical reproduction of methods that are useful in experimental sciences might even prove counterproductive. Indeed, whether or not it is possible to make observations by carrying out controlled experiments is a fundamental factor in determining whether uncertainty is inherent to a particular field of study.

Natural sciences can repeat an experiment up to the point where they obtain relations in which the random component is reduced to a bare minimum, which would enable them to use a quasi-deterministic rather than a probabilistic language. However, non-experimental sciences usually have to deal with a context of great uncertainty and, in consequence, need a probabilistic language.
Sims (1996) recently underscored the factors which make economics a singular science: data scarcity, the existence of alternative explanatory theories, and the pressure for results exercised by decision-makers. Were any one of these factors to be eliminated, it would not be strictly necessary to use probabilistic language: if data were abundant we could more accurately discriminate between alternative theories. If there were only a single theory, discrimination would be unnecessary. It would also be unnecessary if the effects of the decisions made were independent of the theory chosen. Unfortunately, taken together, these three factors render the probabilistic nature of economic evidence unavoidable.

Moreover, it seems evident that once uncertainty is accepted as an inherent part of the learning process in economics, it should be treated with the rigour provided by statistical inference when applied to probabilistic models. In other words, it is not enough to consider a model sufficiently tested when it is capable of generating certain empirical regularities which are “similar” to those observed. If the aim is to discriminate between alternative models then the meaning of “similar” must be specified and that makes it necessary to define goodness of fit criteria, thereby moving into the field of statistical inference. Descriptive statistics is not enough to guarantee the basic principle of modern scientific discourse, according to which theory should match evidence.

Thus we might ask what type of inference is most suitable for economics: Classical or Bayesian. Classical methods reflect the specificity of the experimental sciences for which they were originally developed. As mentioned earlier, in these sciences it may be feasible to control the level of uncertainty and reduce it to a error term. In this case it is reasonable to treat the coefficients of a model as parameters, just as one does with classical inference. However, in the field of economics the level of uncertainty is high and no distinction can be made between the uncertainty that affects the co-efficients themselves and the uncertainty that affects other stochastic elements of the model. In a situation like this, logic suggests that uncertainty as regards the model itself and
uncertainty as it affects the model's components should be treated symmetrically, giving the whole a stochastic configuration about which data allow us to learn. This is Bayesian inference, and it therefore seems natural to use it in economics.

Nowadays it is highly likely that two econometricians will analyse the same sample information and arrive at different conclusions. This signals a lack of discipline which reduces the credibility of econometric analysis. The development of strategies to make specification uncertainty explicit in the modelling process would provide a disciplinary tool for empirical analysis in economics. It would allow to systematise such analysis and make it able to generate shared, widely accepted, and therefore objective relationships just as physical experiments generate widely accepted relationships in experimental sciences. This could be a promising path which might eventually lead to a widely accepted paradigm in the field of empirical macroeconomics, similar to what the general equilibrium paradigm represents in the field of macroeconomic theory.

References


