CESifo Working Papers

www.cesifo.org/wp

Beyond the DSGE Straitjacket

M. Hashem Pesaran
Ron P. Smith

CESifo Working Paper No. 3447
Category 12: Empirical and Theoretical Methods
May 2011

CESifo *Working Paper No. 3447*

Beyond the DSGE Straitjacket

Abstract

Academic macroeconomics and the research department of central banks have come to be dominated by Dynamic, Stochastic, General Equilibrium (DSGE) models based on micro-foundations of optimising representative agents with rational expectations. We argue that the dominance of this particular sort of DSGE and the resistance of some in the profession to alternatives has become a straitjacket that restricts empirical and theoretical experimentation and inhibits innovation and that the profession should embrace a more flexible approach to macroeconometric modelling. We describe one possible approach.

JEL-Code: C100, E100.

Keywords: macroeconometric models, DSGE, VARs, long run theory.

M. Hashem Pesaran
Faculty of Economics
Cambridge University
UK – Cambridge, CB3 9DD
mhp1@econ.cam.ac.uk

Ron P. Smith
Department of Economics
Birkbeck College / University of London
UK - London WC1E 7HX
r.smith@bbk.ac.uk

13 April 2011
This paper is based on a presentation by the first author at the special session oo the “Future of Macroeconomics” held at the MMF 2010 Conference, Cyprus, 29 August to 5 September, 2010.

1 Introduction

Academic macroeconomics and the research department of central banks have come to be dominated by Dynamic Stochastic General Equilibrium (DSGE) models based on micro-foundations of representative agents, who solve inter-temporal optimisation problems under rational expectations. New Keynesian variants will also allow for some frictions in adjustment. Typical examples are Smets and Wouters (2003, 2007) and Christiano et al. (2005). Schorfheide (2011) provides a recent review of the approach. There are some who have criticised the new Keynesian DSGE models, for having sacrificed their theoretical coherence in an attempt to fit the data. Chari et al. (2009), among others, argue that theoretical consistency, particularly within an optimising framework, is essential for policy relevance, and that the relevant data to fit is microeconomic data, not 'the same old macroeconomic data'. They criticise new Keynesian models like Smets and Wouters, for adding free parameters to fit the data, and doubt that the shocks are actually structural or consistent with micro data. They say, p243, 'The tradition favoured by many neoclassicals (including us) is to keep a macro model simple, keep the number of its parameters small and well motivated by micro facts, and put up with the reality that no model can, or should, fit most aspects of the data.' In this neoclassical tradition, the number of parameters is kept small by using very special functional forms, such as power utility functions and Cobb-Douglas production functions and their treatment of micro facts clashes with mainstream microeconometric practice, which rather than keeping the number of parameters small, has increasingly adopted semi-parameteric or non-parametric approaches. Aggregation across heterogeneous micro decision rules invariably lead to macro relations with very different dynamic properties, and no simple extrapolation from micro to macro behaviour seem possible.

Chari et al. criticise new Keynesian DSGE models for not being close enough to the neoclassical tradition, whilst on the contrary we argue that effective macroeconometric modelling requires a sensitive trade-off between consistency with theory, adequacy in representing the macroeconomic data, and relevance to a particular purpose, in the case of macroeconomic models the main purpose is to inform macroeconomic policy.[1] In the light of these three criteria, we would argue that while the DSGE approach has provided many theoretical insights, the dominance of this approach and the resistance of some macroeconomists to alternatives has become a straitjacket that restricts empirical and theoretical experimentation, inhibits innovation and reduces the contribution economists can make to policy debates. The DSGE insistence on a particular type of micro-founded theory, when there are other types of theory available, has been at the expense of adequately representing the data and of being relevant to central policy issues. The type of questions that the DSGE approach focuses on, such as how to identify shocks, are often not the most relevant questions, for understanding the economy or informing policy. But this narrow focus has led to the neglect of wider questions, where theory is less developed, there are more puzzles and the DSGE approach is outside of its comfort zone.

We are not alone in raising concerns about the dominance of the DSGE approach in macro modelling. Since the 2007 financial crisis and the subsequent 'great recession', a number of dissenting views have appeared in the literature. For example, following Hayek,

[1] These three criteria are discussed in more detail in Pesaran & Smith (1985).

Caballero (2010) discusses the ‘pretence-of-knowledge’ syndrome and argues that the DSGE approach confuses the precision it achieved within its own narrowly defined framework with the precision it achieved about the real world. He also argues that we should be in ‘broad exploration’ mode, with much more diversification of research and methodology than is currently accepted. De Grauwe (2010) says: ‘DSGE models provide a coherent framework of analysis. This coherence is brought about by restricting acceptable behaviour of agents to dynamic utility maximisation and rational expectations. The problem of DSGE models (and more generally of macroeconomic models based on rational expectations) is that they attribute extraordinary cognitive capabilities to individual agents. In addition, these models rely on a number of *ad hoc* assumptions about habit persistence, search and adjustments costs to fit the data.’ He emphasises the endogenous inertia that comes from private sector information acquisition and learning.

There has also been much comment on the fact that the DSGE models used by central banks typically did not and some still do not include the foreign, financial and housing variables that proved crucial to the transmission of the crisis and also have very limited treatment of the government or fiscal policy. These omissions made the models less relevant to central policy issues. Although there are large specialist literatures on all these issues, these dimensions were treated as peripheral and not integrated into the core of macroeconomics, where the literature was dominated by a narrower DSGE approach. These other dimensions were either ignored, or the models were adapted in such a way as to maintain the core real business cycle, RBC, structure.

We first discuss crucial differences between microeconomics and macroeconomics, then review the DSGE approach and consider a number of questions about the orthodox DSGE methodology with respect to theoretical assumptions, identification and estimation. In the light of this discussion, like Caballero (2010), we suggest that it is more useful to adopt a less dogmatic, more flexible approach to what constitutes acceptable theory and methodology, treating the theory as a guide to empirical work that may inform policy rather than a straitjacket. One benefit of this more flexible agnostic approach is that it avoids what Manski (2010) calls ‘policy analysis with incredible certitude’.

A more flexible approach would involve considering a wider range of data, including global and financial variables and tolerating multiple models, within which various types of theory are interpreted flexibly. The theory would provide lists of variables that interact strongly, crucial constraints of the sort that determine long-run relations, some restrictions on the signs of the interaction between variables and some guidance on functional forms. It also seems likely that, contrary to the DSGE approach, the theory is more informative about the long-run, where arbitrage conditions are important, than about transitory dynamics, where the data should get greater weight, see Pesaran, 1997. This is consistent with a long tradition in economics of being agnostic about short-run dynamics, including Milton Friedman’s emphasis on long and variable lags.

2 Elements of DSGE modelling

2.1 The relevance of micro foundations to macro policy

Theoretically consistent micro-foundations, based on either a representative agent or a continuum of agents that can be aggregated, have been seen as a necessary condition for acceptable macroeconometric modelling. While micro-foundations, if available, may be useful, regarding them as the defining quality of an acceptable macro model fails to recognise the difference between microeconomics and macroeconomics. A central criteria for evaluation of macroeconomic models must be their relevance for government decisions about fiscal, monetary and financial stability policy. There may be some who insist that the government should have little or no role in macro management apart from price stability, but given the extent of government expenditure, taxation and regulation, particularly in the financial sphere, it is essential that the models should inform macroeconomic policy, otherwise policy will be swayed by relatively uninformed commentators, usually with vested interests.

We agree that for policy one needs a structural, rather than a purely statistical model, one whose parameters do not change substantially when policy changes. This raises two issues: whether parameters are stable over time, and whether any instability is a response to a policy intervention. While there is considerable evidence for structural breaks in the coefficients and variances of economic models, which must make us cautious in making quantitative policy recommendations; there is little evidence that these breaks can be attributed solely to policy interventions, the central point of the Lucas critique. The usual argument by DSGE proponents is that their parameters are structural because they depend on fundamental microeconomic aspects of preferences and technology. But the DSGE models are being used to describe the macroeconomic behaviour of the aggregate economy and it has long been known that the conditions necessary to aggregate the optimising decision rules of individual agents to an aggregate relationship of the same form are rather stringent. Aggregate relationships will not be of the same form as those for an individual agent and will typically involve other features of the distribution of the micro variables, than just averages. Priors from microeconometric work will not be appropriate for what appears to be their macroeconomic equivalent because the context is different, a point made by Hansen and Heckman (1996). General equilibrium involves the interaction of many heterogenous individual agents, subject to correlated shocks, where learning and markets may be incomplete. Even if the micro-economic parameters were structural, aggregation over interacting, heterogeneous agents may make their macro analogues non-structural. There has been some concern within the DSGE approach about how structural the parameters are. Fernandez-Villaverde and Rubio-Ramirez (2007) provide evidence that the structural parameters change within the sample. Cogley and Yagihashi (2010) examine the effect of misspecification arising from approximation. Chang, Kim and Schorfheide (2010) find that taking account of aggregation and features of shock distributions can have a major effect on the properties of the model.

With heterogeneous individual agents, it is assumed that the random errors of the individuals cancel out across agents. This requires that the individual errors to be cross-sectionally independent or at least only weakly correlated. Pesaran and Chudik (2011) consider the problem of aggregation where each micro unit is potentially related to all other micro units and where shocks to each micro unit can be correlated across units both contemporaneously and with a lag. In such circumstances, the dynamic time-series properties of the aggregate variables can be fundamentally different from those of the underlying micro units. They illustrate this issue with disaggregated data on European consumer prices, where it is known that while prices at the micro level are quite flexible, aggregate inflation tends to be quite persistent. This aggregate persistence results from the combination of cross section heterogeneity with persistence in the unobserved common factors that drive the shocks. This then argues for more flexible dynamics in the model for the aggregate than would be appropriate in the model for the micro units.

Policy involves setting some decision or control variables, the policy instruments, subject to certain constraints, for a given choice of the conditional probability distribution of the state variables, in order to maximise some suitable objective function. Whereas the objective functions of firms and consumers, profits and utility, are relatively uncontroversial, the objective function of governments is more controversial. The controversies arise from the difficulty of establishing a social welfare function (which will reflect considerations of both efficiency and distribution) and the contribution of political economy factors to the determination and implementation of policy. Whereas individual consumers and firms can, in most cases, treat the probability distribution of the state variables as invariant to their actions, governments certainly cannot. Because they are such a large actor in the economy, government decisions will impact on the economy. In addition, there are inevitably strategic aspects, such as general perceptions of the credibility of the policy. The information available to the government and other agents may well be different, perhaps because the government is strategic in how it reports economic data.

The constraints involved may be technological, financial, legal or institutional; may be quite difficult to specify *ex ante* and may be subject to change. For instance, many cases of fiscal innovation, which allow governments to increase tax rates substantially, have been in time of war, and wars play a major role in political economy theories of the evolution of state capacity, particularly fiscal capacity, Besley and Persson (2010). Specification of decision variables may be problematic. For instance, in monetary policy the decision variable was once seen as the money supply, then seen as interest rates, until the zero lower bound constraint became binding, when the decision variable become large scale asset purchases (commonly known as quantitative easing). All these uncertainties should be included in the decision problem for robust control, but doing so raises difficulties. Caballero (2010) comments 'Academic models often provide precise policy prescriptions because the structures, states, and mechanisms are sharply defined. In contrast, policy makers do not have these luxuries.'

The policy maker then has to solve this decision problem in 'real time' over some future horizon, and this may be difficult to compute and implement in a timely manner. Some decisions are also made by committees which further complicates the decision process. It is usually assumed that unique stable solutions to such decision problems exist, which they may well not. It is also assumed that if a solution does exist it can be implemented. But operational obstacles to the implementation of a policy may only become apparent when one attempts to apply the solution.

For the policy maker, the macro-economy may be characterised as a complex system, with many interacting heterogeneous agents operating in a non-stationary environment with various weak and strong links between them, subject to occasional contagion and herding effects and other structural breaks. The usual micro-foundations of standard DSGE models may not capture such complexity. There are also a range of features of the data, such as the observed persistence of the variables, the equity premium puzzle and the forward premium puzzle, that raise fundamental questions about the basic framework. Macro DSGE models tended to respond to these features either by introducing *ad hoc* frictions that do not disturb the basic structure or by restricting the observed variables included in the model to avoid confronting the puzzles directly, thus creating a protective belt of procedures around the core, that insulate them from falsification.

2.2 Optimisation by representative agents with rational expectations

The rational expectations hypothesis, REH, has been extensively criticised, by for instance Pesaran (1987), Manski (2004), Hansen and Sargent (2008), Fuster et al. (2010). The critics have emphasised a range of problems including the amount of knowledge required, the problems of learning, doubts about the true model, ambiguity aversion, behavioural uncertainty where one must form expectations about the expectations of others and the limitations of the linear-quadratic model needed for certainty equivalence. Some of these difficulties and some of the identification problems discussed below can be dealt with by more extensive use of survey measures of expectations. However, inter-temporal optimisation calculations typically require expectations far into the future and survey measures of distant expectations are rarely available. Therefore, even with survey measures one would need to model the expectations formation process in order to provide estimates of these more distant expectations.

According to the REH the subjective characterisation of uncertainty as conditional probability distributions will coincide with the associated objective probability outcomes. The REH is mathematically elegant, allows model consistent solutions and fits nicely with equilibrium theory. Although less demanding than perfect foresight the REH requires economic agents to know or learn the 'true' conditional probability distributions. This assumption is particularly problematic with behavioural uncertainty, where agents need to form expectations about the expectations of others, as in Keynes's beauty contest. While the REH may be a reasonable working hypothesis for processes that are stationary and ergodic, many economic and financial processes are continually affected by institutional, technological and political changes which are largely unpredictable and show up as structural breaks in the parameters of estimated relationships. Expected present value relationships, which play a key role in economics (for the permanent income theory) and finance (where they provide fundamentals), require long-horizon predictions from a finite available history. This implies that either the present values cannot be calculated under parameter uncertainty, for instance about future structural breaks, or that estimated present values do not have moments and are therefore not a reliable guide to the future, Pesaran, Pettenuzzo and Timmerman (2007).

It is usually assumed that the agents face complete markets, which clear efficiently and the frictions introduced into DSGE models have typically been in price and wage adjustment, not the more fundamental frictions that have been emphasised, for instance, in the search and learning literatures. The absence of current (e.g. contracts for second hand capital goods) and contingent (contracts conditional on the agent being liquidity constrained) markets means that the agents condition on unobservable (to the econometrician) shadow prices. Pesaran & Smith (1995) discuss the treatment of such shadow prices.

In DSGE models a lot of markets tend to be neglected. Most DSGE models are closed economy, though there have been some extensions to two block structures and small open economy versions. But the general impression is that one sees globalisation everywhere except in the DSGE models. The DSGE theory naturally extends to financial variables, since the consumption Euler equation which is the centre of the DSGE model is also the basis for a large amount of finance theory. However, the DSGE models have typically been restricted not to cover financial variables such as long interest rates, equity prices and exchange rates. A possible reason for this restriction is a desire to keep the model simple and to avoid confronting puzzles like the forward premium and equity premium puzzle, that cast doubt on the basic framework.

2.3 Identification and estimation

Most DSGEs are constructed by linearising underlying non-linear models around their steady state.[2] The parameters of the linearised solution will be non-linear functions of the underlying structural parameters, say $\boldsymbol{\theta}$. The structural parameters of DSGE models tend to include micro-economic parameters like the discount rate or the coefficient of risk aversion; policy parameters such as the relative weight put on inflation and unemployment in the policy loss function; and the coefficients and variances of autoregressive shocks. The objects of interest are often the structural shocks and the impulse response functions of the system to those shocks. To measure these one needs to identify and estimate $\boldsymbol{\theta}$.

To organise the discussion, consider a linearised rational expectations model for an $n \times 1$ vector of observed variables of interest, $\widetilde{\mathbf{y}}_t$, $t = 1, 2, ..., T$ using the tilde to denote that they are measured as deviations from their steady state, a $k \times 1$ vector of other, observed and unobserved, variables of less direct interest, $\widetilde{\mathbf{w}}_t$, (which may contain further lags) again measured as deviations from steady state, and an $s \times 1$ vector of mean zero serially uncorrelated structural shocks, $\boldsymbol{\varepsilon}_t$, with $E(\varepsilon_t \varepsilon_t') = \Omega(\boldsymbol{\theta})$. It is common to assume that $\Omega(\boldsymbol{\theta}) = \mathbf{I}_n$. Define $\mathbf{z}_t = (\widetilde{\mathbf{y}}_t', \widetilde{\mathbf{w}}_t')'$ then a typical DSGE model can be written as

$$\begin{aligned} \mathbf{B}_0(\boldsymbol{\theta})\mathbf{z}_t &= \mathbf{B}_1(\boldsymbol{\theta})\mathbf{z}_{t-1} + \mathbf{B}_2(\boldsymbol{\theta})E_{t-1}(\mathbf{z}_{t+1}) + \boldsymbol{\varepsilon}_t, \qquad (1) \\ \mathbf{z}_t &= \mathbf{B}_0(\boldsymbol{\theta})^{-1}\mathbf{B}_1(\boldsymbol{\theta})\mathbf{z}_{t-1} + \mathbf{B}_0(\boldsymbol{\theta})^{-1}\mathbf{B}_2(\boldsymbol{\theta})E_{t-1}(\mathbf{z}_{t+1}) + \mathbf{B}_0(\boldsymbol{\theta})^{-1}\boldsymbol{\varepsilon}_t \\ \mathbf{z}_t &= \mathbf{C}_1(\boldsymbol{\theta})\mathbf{z}_{t-1} + \mathbf{C}_2(\boldsymbol{\theta})E_{t-1}(\mathbf{z}_{t+1}) + \mathbf{R}(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t \end{aligned}$$

where the linear coefficient matrices are possibly non-linear functions of a vector of deeper parameters, $\boldsymbol{\theta}$, some of which may be thought to be more structural, or stable. $E_{t-1}(\mathbf{z}_{t+1}) = E(\mathbf{z}_{t+1} \mid \mathcal{I}_{t-1})$ and $\mathcal{I}_{t-1}$ is the common information set available at time $t-1$. The nature of the solution then depends on the roots of the quadratic matrix equation

$$\mathbf{C}_1(\boldsymbol{\theta})\boldsymbol{\Phi}^2 - \boldsymbol{\Phi} + \mathbf{C}_2(\boldsymbol{\theta}) = 0. \qquad (2)$$

There will be a unique stationary solution if (2) has a real matrix solution such that all the eigenvalues of $\boldsymbol{\Phi}$ and $(\mathbf{I} - \mathbf{C}_1(\boldsymbol{\theta})\boldsymbol{\Phi})^{-1}$, lie strictly inside the unit circle. The solution, if

[2] Higher order approximations are starting to be used.

it exists, takes the form $\mathbf{z}_t = \mathbf{A}(\boldsymbol{\theta})\mathbf{z}_{t-1} + \mathbf{R}(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t$, which can be expressed in terms of the variables of interest as

$$\begin{aligned}\widetilde{\mathbf{y}}_t &= \mathbf{A}_{11}(\boldsymbol{\theta})\widetilde{\mathbf{y}}_{t-1} + \mathbf{A}_{12}(\boldsymbol{\theta})\widetilde{\mathbf{w}}_{t-1} + \mathbf{R}_1(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t \\ \widetilde{\mathbf{w}}_t &= \mathbf{A}_{21}(\boldsymbol{\theta})\widetilde{\mathbf{y}}_{t-1} + \mathbf{A}_{22}(\boldsymbol{\theta})\widetilde{\mathbf{w}}_{t-1} + \mathbf{R}_2(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t.\end{aligned}$$

Within this framework one can consider the theoretical motivation for the particular structures used, the nature of the solution and the identification and estimation of $\boldsymbol{\theta}$.

This framework is quite sensitive to quite minor forms of model misspecification. Suppose, that in forming their expectations, the agents know the form of the solution but not the exact non-linear mapping from the solution parameters to $\boldsymbol{\theta}$. Then the cross-equation restrictions implied by θ do not hold. Even this weak form of bounded rationality invalidates the relation of the model to the deeper parameters.

To make this specific suppose that there are no unobservables in (1) and that the agents know $\mathbf{B}_0\widetilde{\mathbf{y}}_t = \mathbf{B}_2\widetilde{\mathbf{y}}_{t-1}+\mathbf{B}_1 E_{t-1}(\widetilde{\mathbf{y}}_{t+1})+\boldsymbol{\varepsilon}_t$, and that the solution is of the form $\widetilde{\mathbf{y}}_t = \mathbf{A}\widetilde{\mathbf{y}}_{t-1}+\mathbf{v}_t$. Using some estimate of A their expectation is given by $E_{t-1}(\widetilde{\mathbf{y}}_{t+1}) = \mathbf{A}^2\widetilde{\mathbf{y}}_{t-1}$, thus the data generating process becomes

$$\mathbf{B}_0\widetilde{\mathbf{y}}_t = (\mathbf{B}_2 + \mathbf{B}_1\mathbf{A}^2)\widetilde{\mathbf{y}}_{t-1} + \boldsymbol{\varepsilon}_t,$$

without the dependence on $\boldsymbol{\theta}$.

The identification and estimation of $\boldsymbol{\theta}$ is typically done by Bayesian methods using the Kalman Filter to deal with any unobserved elements of $\widetilde{\mathbf{w}}_t$. The Bayesian approach is natural in this context since the estimated DSGE models developed from calibrated RBC models, thus the calibrated parameters provided natural prior means. In early models the number of shocks was less than the number of observed variables, implying exact linear dependencies amongst the variables. Since this is not observed in the data, measurement errors were introduced to account for the lack of linear dependence, which raises questions about the identification of the shocks. Later models had either the same number, or more, shocks than variables.

In the early days of RE, there was considerable interest in issues of observational equivalence, Sargent (1976), and identification, e.g. Wallis (1980), Pesaran (1981), in RE models. However, as the focus shifted from estimation to calibration, interest waned. With the revival of estimation of DSGE models, there has been increased concern about the identification of the parameters, for instance, Canova and Sala (2009) and others have argued that many parameters of these models are poorly identified. Identification in Bayesian DSGE models, is considered in more detail in Koop, Pesaran and Smith (2011), KPS. It is common in these models to judge identification by considering whether prior and posterior differ, but prior and posterior will differ in many cases even if a parameter is unidentified and the priors assumed between the unidentified and identified structural parameters are independent. KPS suggest two Bayesian identification indicators that do not suffer from this difficulty and are relatively easy to compute. The first applies to DSGE models where the parameters can be partitioned into those that are known to be identified and the rest where it is not known whether they are identified. In such cases the marginal posterior of an unidentified parameter will equal the posterior expectation of the prior for that parameter conditional on the identified parameters. The second indicator is more generally applicable and considers the rate at which the posterior precision (inverse of the variance) gets updated as the sample size (T) is increased. For identified parameters the posterior precision rises with T, whilst for an unidentified parameter its posterior precision may increase but its rate of update will be slower than T. This result assumes that the identified parameters are $\sqrt{T}$-consistent, but similar differential rates of updates for identified and unidentified parameters can be established in the case of super consistent estimators. KPS illustrate these results by means of simple DSGE models.

3 Avoiding the straitjacket

As noted in the introduction, general equilibrium emphasises the inter-connections of economic activities, through a range of feedback and network effects and the behaviour of an economy as a system which generally differs from the behaviour of the individuals who comprise it. Representative agent models lose both the complexity of the interconnections between agents and the compositional effects which come from aggregation of the behaviour of the interacting individual agents into a system.

We have argued that theory, while essential, should be regarded as a flexible framework rather than a straitjacket, because features that the theory abstracts from may be important in practice. These features include issues such as aggregation and dynamics, as well as the various puzzles for the standard theory. Even if there is no generally accepted theoretical framework it is important to allow for these features and linkages, at least through more flexible empirical models which allow for quantitative interactions between variables. It is important to distinguish the cases where theory can be used to provide qualitative insights, which may inform judgement, from the case where policy requires a quantitative model. It is also important to distinguish the questions that require an identified model and the questions for which a reduced form model may perform adequately.

In the more flexible approach we would emphasise the following five elements. Firstly, the use of long-run cointegrating relations where they exist. Most DSGE models are log-linearised around a long-run steady state. The steady states are estimated by a constant and perhaps deterministic trend or by a purely statistical method, like the univariate Hodrik-Prescott, HP, filter. The use of statistical filters rather than an economic model for the steady state, indicates low degree of belief in long-run economic relations, and a high degree of belief in the short-run relations to which the DSGE theory applies. This may be the wrong way round, it might be better to use the theory, including any long-run cointegrating relations, to get the steady state and leave the short-run dynamics less restricted. Use of any cointegrating relations found in the data gives the model transparent long-run properties that can be evaluated empirically and can often be given a theoretical interpretation in terms of arbitrage or inter-temporal solvency conditions. There may also be more homogeneity in the long-run relations than in the short-run dynamics.

Secondly, the use of more flexible short-run dynamics. As noted earlier the macroeconomic outcome produced by heterogeneous agents subject to possibly correlated shocks may not correspond to the behaviour of a representative agent and there are many other mechanisms that may induce slow adjustment including habits, expectational errors, learning, and the costs and frictions of search and matching. It may be difficult to model all these mechanisms in a rigorous way, thus it may be sensible to approximate them by longer lags. Thus while theoretical structures like DSGE are very useful for motivating models they should not constrain them too tightly, particularly when the restrictions are strongly rejected by the data. This is also associated with relaxing the restrictions of rational expectations and may involve greater use of survey measures of expectations, as discussed above. Estimation of the long-run relations with flexible short-run dynamics can be done within the context of cointegrating VAR structures, which can also be used to estimate the steady states, as long-horizon forecasts. One can then use the deviations from the steady states to estimate more structural models with identified shocks, as is done in Dees et al. (2009, 2010). Dees et al. (2010) use the deviations from steady state measured in this way to estimate and solve for 33 countries a standard New Keynesian rational expectations model of Phillips Curve, Taylor Rule and IS curve, which includes foreign variables, augmented by a real effective exchange rate equation.

Thirdly, the recognition of the wide range of inter-connections between heterogeneous agents that exist within any economic system. In a macro-economic context this requires a multi-country system, which includes trade and financial variables that provide the main channels of transmission of information and shocks across economies, such as that used in the Global VAR, GVAR, of Dees et al. (2007).

Fourthly, the wide range of inter-connections in the economic system poses issues of dimensionality, since there are inevitably going to be a large number of variables and decision making units involved. Therefore, procedures that use theory and the structure of the data to overcome the curse of dimensionality are required. An example is the infinite VAR, IVAR structure for high dimensional inter-connected systems of Chudik & Pesaran (2011a,b). Central to the analysis of such systems is the role of weak (neighbourhood) and strong (common factor) dependence and the role of a dominant unit, if any. A dominant unit, such as the US in the world economy, influences the rest of the variables in the IVAR, both directly and indirectly, introducing infinite order distributed lags. Nonetheless the effect of the dominant unit, as well as neighbourhood units, can be consistently estimated. The GVAR is a particular application of the IVAR approach.

Fifthly, the wide range of interconnections raises questions about the treatment of shocks. Multi-country VARs, including the GVAR, above have been used to model the international transmission of shocks and there are a range of interesting questions about the transmission of shocks, e.g. about the effect of oil price shocks and the rise of China on the world economy, that can be addressed without a precise identification of the nature of the shock. But in other cases, it has been difficult to give such shocks a clear economic interpretation. The issue of shock identification is addressed in Dees et al. (2010) who provide estimates of the effects of identified shocks for 33 countries using a multi-country version of the familiar rational expectations New-Keynesian (NK) model. The Phillips curve error is interpreted as a supply shock, the Taylor rule error as a monetary policy shock and the IS curve error as a demand shock. As usual, these structural shocks are assumed uncorrelated within a country, but need not be uncorrelated across countries. Supply shocks from different countries may be correlated, since technology is transferred between countries, as may demand, monetary policy, and exchange rate shocks. Thus there are international linkages both directly, through foreign variables in the equations, and indirectly, through error spillover effects.

We now provide more detail on these elements. Firstly with respect to the use of the long-run and more flexible short-run dynamics. Since many growth models have a log-linear representation, suppose that there exists a log-linear approximation to the system, which can be represented as a cointegrating VECM

$$\Delta \mathbf{y}_t = \mathbf{c} - \boldsymbol{\alpha}\boldsymbol{\beta}' \left[\mathbf{y}_{t-1} - \boldsymbol{\gamma}(t-1)\right] + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \mathbf{y}_{t-i} + \mathbf{u}_t, \tag{3}$$

where γ is a $n \times 1$ vector of fixed constants, α is the $n \times r$ matrix of the loading coefficients, and r is the cointegrating rank and β is the $n \times r$ cointegrating matrix, and c is a $n \times 1$ vector of fixed constants. We discuss applying this to high dimensional systems below. The dynamics of unrestricted VECMs are often difficult to interpret, but they allow one to identify the long-run cointegrating relationships.

To derive the steady states or permanent components, write the VECM, (3), as the VAR(p) specification

$$\mathbf{y}_t = \mathbf{b}_0 + \mathbf{b}_1 t + \sum_{i=1}^{p} \mathbf{F}_i \mathbf{y}_{t-i} + \mathbf{u}_t, \tag{4}$$

where $b_0 = (c - \alpha\beta'\gamma)$, $b_1 = \alpha\beta'\gamma$, and $F_1 = (I_n + \Gamma_1 - \alpha\beta^{'})$, $F_i = (\Gamma_i - \Gamma_{i-1})$, $i = 2, ..., p-1$, $F_p = -\Gamma_{p-1}$.Using (4) we can now write down the solution of y_t as

$$\mathbf{y}_t = \boldsymbol{\mu} + \mathbf{g}t + \mathbf{C}\left(1\right)\mathbf{S}_t + \mathbf{C}^*\left(L\right)\mathbf{u}_t, \tag{5}$$

where $\mu = x_0 - C^*(L)u_0$, $S_t = \sum_{j=1}^{t} u_j$, $C^*\left(L\right) = \sum_{j=0}^{\infty} C_j^* L^j$,

$$\mathbf{C}_j = \mathbf{C}_{j-1}\mathbf{F}_1 + \mathbf{C}_{j-2}\mathbf{F}_2 + \cdots + \mathbf{C}_{j-p}\mathbf{F}_p, \text{ for } j = 1, 2, ...,$$

$C_0 = I_k$, $C_1 = -(I_k - F_1)$, and $C_j = 0$ for $j < 0$; $C_j^* = C_{j-1}^* + C_j$, for $j = 1, 2, ...$, with $C_0^* = C_0 - C(1)$, and $C(1) = \sum_{j=0}^{\infty} C_j$. Hence, it is easily seen that the steady state or long-horizon expectation is

$$\mathbf{y}_{st}^P = \lim_{h \to \infty} E_t\left[\mathbf{y}_{t+h} - \boldsymbol{\mu} - \mathbf{g}(t+h)\right] = \mathbf{C}\left(1\right)\sum_{j=1}^{t} \mathbf{u}_j = \mathbf{C}\left(1\right)\mathbf{S}_t, \tag{6}$$

which is the multivariate version of the Beveridge-Nelson (BN) stochastic trend component. Note that y_{st}^P is uniquely determined from the time series observations on x_t and its lagged values.

Dees et al. (2009, 2010) estimate an unrestricted cointegrating GVAR and then estimate the steady state, or permanent components, as the long horizon forecasts from this cointegrating GVAR. Denote these first round estimates as y_t^P. Using these estimates, the deviations from steady state are then calculated as :

$$\widetilde{\mathbf{y}}_t = \mathbf{y}_t - \mathbf{y}_t^P = \mathbf{y}_t - \left(\boldsymbol{y}_0 + \boldsymbol{\delta}t + \mathbf{C}(1)\mathbf{S}_t\right)$$

where S_t is a vector of stochastic trends. Treating the steady states as constants or deterministic trends, sets $C(1) = 0$ and incorrectly ignores these stochastic trends.

These deviations can be used to estimate rational expectations, DSGE type, model like (1) where one can either impose the restriction implied by θ or not. Given that one has estimated the forward looking DSGE model, one can obtain a solution computed as

$$\widetilde{\mathbf{y}}_t = \Phi(\boldsymbol{\theta})\widetilde{\mathbf{y}}_{t-1} + \mathbf{R}(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t \tag{7}$$

where ε_t are the structural errors rather than the errors from the cointegrating VAR u_t. Since the deviations from steady state, $\widetilde{\mathbf{y}}_t$ are by definition stationary all the roots of $\Phi(\theta)$ must lie within the unit circle. Notice that a problem with the HP filter is that it does not guarantee that $\widetilde{\mathbf{y}}_t$ will be stationary. Using the fact that the roots of $\Phi(\theta)$ lie within the unit circle, we can rewrite (7) as $y_t - y_t^P = \Phi(\theta)(y_{t-1} - y_{t-1}^P) + R(\theta)\varepsilon_t$, and hence

$$\mathbf{y}_t = \mathbf{y}_t^P + \left[\mathbf{I}_n - \boldsymbol{\Phi}(\boldsymbol{\theta})L\right]^{-1}\mathbf{R}(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t.$$

Given that $\Phi(\theta)$ and $R(\theta)$ have been estimated then one could get second round estimates of the permanent components

$$\widehat{\mathbf{y}}_t^P = \mathbf{y}_t - \left[\mathbf{I}_n - \boldsymbol{\Phi}(\hat{\boldsymbol{\theta}})L\right]^{-1}\mathbf{R}(\hat{\boldsymbol{\theta}})\boldsymbol{\varepsilon}_t.$$

These will be more efficient, if the parametric restrictions embodied in θ hold. But given the danger of these second round estimates being biased by model misspecification, perhaps of the sort discussed above where agents did not know the mapping of $\Phi(\theta)$ to θ, it seems safer to rely on the first round estimates of the permanent components. Dees et al. (2010) use the deviations from steady state measured in this way to estimate and solve for 33 countries a standard New Keynesian rational expectations model of Phillips Curve, Taylor Rule and IS curve, which includes foreign variables, augmented by a real effective exchange rate equation.

With respect to interconnections and dimensionality, as was noted above a major issue with the current stock of DSGE models is lack of global and financial interlinkages. Given the degree of inter-connection in the world economy, the set of potentially important variables is very large and there is a choice as to how many different countries, different sectors (real, financial, housing) are included. If one includes many sectors and countries one rapidly runs into the curse of dimensionality problem. To deal with this one either has to shrink the data (e.g. just use observed aggregates or estimate common factors) or shrink the parameter space. In the judgement on how this is done, the role of a dominant unit or factor plays a central role, both for exogeniety in estimation and for the degree of interconnection of the system, as in the IVAR and GVAR approach.

GVAR methodology allows one to decouple a model of world economy, into sub-systems as country specific models which can be consistently estimated separately, but analysed simultaneously taking account of the fact that all the variables in the underlying global model are endogenous and linked. The estimated GVAR strongly indicates the importance of the international and financial linkages. The GVAR structure has also been used for the analysis of a large variety of other high dimension problems.

4 Conclusion

Formal theory is essential in enabling us to organise our *a priori* knowledge about a problem in a coherent and consistent way. But the formal theory must be confronted with the data if it is to enhance our understanding and have relevance to the practical problems of macroeconomic policy. We have argued that macroeconometric modelling would benefit from a more flexible approach which does not require narrow adherence to one particular theoretical framework. In the process one would need to be more explicit about the trade-offs between consistency with theory, adequately representing the data and relevance for particular purposes. This will inevitably involve the use of different models for different purposes. This raises new issues of model evaluation that go beyond purely statistical measures of fit and parsimony.

We have suggested an approach that uses the long-run cointegrating information in the data, but allows more flexible short-run dynamics; recognises the inter-connectedness of large systems and develops methods to estimate high dimensional systems that help identify certain types of shocks. A flexible approach to theory could also make the task of model evaluation and comparison much more difficult and is likely to lead to another layer of uncertainty about policy and about 'good' academic practice. Some may find this lack of agreement and uncertainty psychologically disturbing, but if the previously agreed framework was wrong and the certainty about appropriate policy unjustified, it seems an improvement.

References

Besley, T. & T. Persson (2010) State capacity, conflict and development Econometrica 78(1) 1-34.

Caballero, R.J. (2010) Macroeconomics after the crisis: time to deal with the pretense-of-knowledge syndrome. Journal of Economic Perspectives, 24(4), 85-102.

Canova, F. and Sala, L. (2009) Back to square one: identification issues in DSGE models, Journal of Monetary Economics, 56, 431-449.

Chang, Y, S-B Kim & F. Schorfheide (2010) Labour-market heterogeneity, aggregation and the Lucas Critique, NBER Working Paper 16401.

Chari, V.V., P.J.Kehoe & E.R. McGrattan (2009) New Keynesian models: not yet useful for policy analysis, American Economic Journal: Macroeconomics, 2009 1(1) 242-266.

Christiano, L.J., M. Eichenbaum, & C.L. Evans (2005) Nominal rigidities and the dynamic effect of a shock to monetary policy, Journal of Political Economy, 113(1), 1-45.

Chudik, A. & M.H. Pesaran (2011a) Infinite dimension VARs and factor models, forthcoming Journal of Econometrics.

Chudik, A. & M.H. Pesaran (2011b) Econometric analysis of high dimensional VARs featuring a dominant unit, Cambridge.

Cogley, T. & T. Yagihashi (2010) Are DSGE approximating models invariant to shifts in policy? The B.E. Journal of Macroeconomics 10(1) article 27.

Dees, S., di Mauro, F., Pesaran, M.H., Smith, L.V. (2007). Exploring the international linkages of the euro area: a global VAR analysis. Journal of Applied Econometrics, 22, 1-38.

Dees, S., M.H. Pesaran, L.V. Smith & R.P. Smith (2009) Identification of new Keynesian Phillips curves from a global perspective, Journal of Money Credit and Banking, 41(7) 1481-1502.

Dees, S., M.H. Pesaran, L.V. Smith and R.P. Smith (2010) Supply, demand and monetary policy shocks in a new Keynesian model, ECB working paper 1239.

De Grauwe, P (2010) The scientific foundations of dynamic stochastic general equilibrium (DSGE) models, Public Choice, 144, 413-443.

Fernandex-Villaverde, J & J.F. Rubio-Ramirez (2007) How structural are structural parameters? NBER Macroeconomics Annual, 83-137.

Fuster, A. D. Laibson & B. Mendel (2010) Natural expectations and macroeconomic fluctuations, Journal of Economic Perspectives, 24(4), 67-84.

Hansen L.P. & J.J. Heckman (1996) The empirical foundations of calibration, Journal of Economic Perspectives, 10(1) 87-104.

Hansen, L.P & T.J. Sargent (2008) Robustness, Princeton University Press.

Koop, G., M.H. Pesaran & R.P. Smith (2011) On identification of Bayesian DSGE models, manuscript.

Manski C. F. (2004) Measuring expectations, Econometrica, 72, 1329-1376.

Manski, C.F. (2010) Policy analysis with incredible certitude, NBER working paper 16207.

Pesaran, M.H. (1981) Identification of Rational Expectations Models, Journal of Econometrics, 16, 375-398.

Pesaran, M.H. (1987) The limits to rational expectations, Blackwell.

Pesaran, M.H. (1997) The role of economic theory in modelling the long-run, Economic Journal, 107, 178-191.

Pesaran, M.H. & A. Chudik (2011) Aggregation in large dynamic panels, Cambridge.

Pesaran, M.H., D. Pettenuzzo and A. Timmerman (2007) Learning, structural instability and present value calculations, Econometric Reviews, 26(2-4) 253-288.

Pesaran, M.H. & R.P. Smith (1985) Evaluation of macroeconometric models, Economic Modelling, 2(2) 125-134.

Pesaran, M.H. & R.P. Smith (1995) The role of theory in econometrics, Journal of Econometrics, 67(1) 61-79.

Sargent T.J. (1976) The observational equivalence of natural and unnatural rate theories of macroeconomics, Journal of Political Economy, 84, 631-40.

Schorfheide, F. (2011) Estimation and evaluation of DSGE models: progress and challenges, NBER working paper 16781.

Smets, F. and R. Wouters (2003) An estimated dynamic stochastic general equilibrium model of the Euro area, Journal of the European Economic Association, 1, 1123-1175.

Smets, F. and R. Wouters (2007) Shocks and frictions in US business cycles: A Bayesian DSGE approach, American Economic Review, 97, 588-606.

Wallis, K. (1980) Econometric implications of the rational expectations hypothesis, Econometrica, 48, 49-73.