DSGE MODELLING FOR MACROECONOMIC POLICIES

(Keynote Address at IGIDR Conference on

DSGE Modelling for Emerging Economies)

(7^{th} August 2018)

INTRODUCTION

Right from the 1970s policymakers have displayed an interest in formal models of the macro-economy with a view to using them for forecasting and policy purposes. Central banks, in particular, have felt the need to take recourse to such models as an aid in the formulation of monetary policy (and in recent years for maintaining financial stability). Typically an array of models is used to throw light on different aspects of policy, while judgment continues to play an important role in the actual policy decisions.

Models in the 1970s : Basically large Simultaneous Equation Models (SEMs) , later followed by multiple time-series models, (MTSMs) which in turn gradually gave way to VARs and Structural VARs in the 1990s.

In the last decade or so an increasing number of central banks are actively engaged in the construction of DSGE (Dynamic Stochastic General Equilibrium) models (e.g. the Bank of England, the Federal Reserve Board, the European Central Bank, the IMF, Sveriges Riksbank etc.).

Most of these banks are in the developed world, but it will not be long before EME central banks follow suit (see Tovar (2008). Or rather, they are doing so already.

ADVANTAGES

DSGE proponents claim at least four major advantages for their model.

1. **Firstly, it is claimed that these models are solidly grounded in economic theory with secure micro-foundations.**
2. **Related to the above, it is maintained that the parameters in the model are structural, and hence invariant to policy shocks. This by-passes the Lucas Critique and enables policy simulations aimed at judging the impacts of policy changes on key macroeconomic variables. This (it is felt) is a major advantage over more data-based traditional models such as VAR or simultaneous equation models.**
3. **DSGE models seem to record a forecasting performance at least comparable to other models (the Bayesian VAR is usually chosen as the benchmark in such comparisons).**
4. **In spite of their elaborate structure, the results of simulations under alternative policy scenarios can be communicated to policymakers in an easily understood manner.**

Several other advantages are also claimed on behalf of the models viz that they bring out the key role of expectations and (being of a general equilibrium nature) can help the policy maker by explicitly highlighting the macro-economic scenarios in response to various contemplated policy outcomes. Additionally, as we discuss later, the models in spite of being strongly tied to theory, can be "taken to the data" (to use a phrase which has become standard in this literature) in a meaningful way. A major feature of these models is that their theoretical underpinnings lie in what has now come to be called as the *New Consensus Macro-economics* (NCM) which established itself in the 1980s as the *weltanschauung* of the bulk of the macroeconomics profession.

BASIC STRUCTURE

The essential ingredients of a typical DSGE model are the following

- **(i) a collection of rational optimizing economic agents including households, final goods producers, intermediate goods producers, labour, government etc., who maximize their separate expected utility functions subject to their respective budget and resource constraints**
- **(ii) the optimization process leads to non-linear Euler equations which need to be solved by special methods such as those given by Blanchard and Kahn (1980), Sims (2002), Uhlig (1999) etc**
- **(iii) the model is then log-linearized around its steady state and**
- **(iv) the system is driven by a number of stochastic shocks, such as shocks to technology, to leisure preference, to government policy rules etc.**

These four features are shared by all DSGE models ranging from the simpler DSGE models like the Real Business cycle model of Campbell (1994), Hansen (1985) etc. to the medium scale models of Christiano et al (2007), Sborodone et al (2010) etc. to the large scale models of Smets and Wouters (2003, 2004), Gali et al (2012), Harrison et al (2005)

STATE-SPACE REPRESENTATION

A surprisingly rich class of DSGE models, irrespective of their scope, can be expressed in a stochastic linear general equilibrium framework and put in the following state space format

$$\begin{bmatrix} E_t(f_{t+1}) \\ s_{t+1} \end{bmatrix} = \mathrm{A} \begin{bmatrix} f_t \\ s_t \end{bmatrix} + \gamma \epsilon_t \qquad \textbf{(1)}$$

where we introduce a distinction between those variables which are predetermined at time t (which includes both exogenous and some endogenous variables) which are termed state variables (denoted by s_t) and those endogenous variables not so predetermined which are termed forward looking, "control" or "jump" variables (denoted by f_t).

This system can be solved provided the conditions mentioned in Blanchard and Kahn (1980), are satisfied. For our further analysis we assume that these conditions are satisfied and write (1) as:

$$f_t = \Gamma s_t \qquad \textbf{(2)}$$

$$s_t = \mathrm{P}_1 s_{t-1} + \mathrm{P}_2 \epsilon_t \qquad \textbf{(3)}$$

where ϵ_t is a vector of shocks[1] and , $\mathbf{P_1}, \mathbf{P_2}$ and Γ are matrices of appropriate dimensions.

We stack the vectors $\begin{bmatrix} f_t \\ s_t \end{bmatrix}$ into a single vector say X_t and rewrite (2)-(3) as:

$$X_t = \mathbf{C}X_{t-1} + \mathbf{B}\epsilon_t \qquad (4)$$

where $\mathbf{C} = \begin{bmatrix} 0 & \Gamma\mathbf{P} \\ 0 & P \end{bmatrix}$ **and** $\mathbf{B} = \begin{bmatrix} \Gamma\mathbf{Q} \\ \mathbf{Q} \end{bmatrix}$

It is tempting to proceed to a direct estimation of the parameters of the model (4). However this fails because most DSGE models suffer from what is called as "the stochastic singularity" problem (see Canova and Sala (2009)).

---

[1] Some elements of ϵ_t could be zero.

STOCHASTIC SINGULARITY

The model is 'stochastically singular' if the spectral density of observed variables $S_{ff}(\lambda)$ is *rank-deficient at almost all frequencies.*

When does it arise?

1. **Number of shocks is less than or equal to number of observables. Theory underpins only few truly structural driving forces. At business cycle frequency, real economic variables are driven by few factors**

2. **A combination of elements in Y is perfectly correlated. Macroeconomic and financial data are never perfectly correlated. But Macroeconomic data feature robust regularities and**

co-movement. Real and nominal cyclical comovement is very strong

Consequences :

1. **It precludes use of likelihood-based methods and Kalman filter. Bayesian-likelihood methods are of no help here, of course**

Overcoming the Singularity Problem :

1. **<u>Calibration :</u> In the older generation of DSGE models, calibration was the generally accepted method of analysis (see e.g. Gregory and Smith (1991), Christiano and Eichenbaum (1992) etc.).[2] Calibration proceeds by attributing certain fixed values to the structural parameters based on economic theory, intuition or more likely previous micro-econometric studies . The model is subsequently simulated using these parameters with a series of artificial shocks. Statistical features of the simulated outcomes (such as unconditional moments or the spectrum) are compared with those of actual data. Appropriate distance measures between the simulated and actual data statistics are computed to arrive at some judgements about the adequacy of the calibrated values of the parameters (see e.g. Watson (1993), DeJong et al (1996) etc.).**
2. **<u>Time-Varying Parameters :</u> This method (usually associated with Smets and Wouters (2003, 2004)) introduces time variation in some of the parameters by subjecting them to stochastic shocks. In practice the number of parameters subject to the shock must be sufficient to overcome the deficit in the number of shocks. Of course as to which parameters are to be treated as fixed and which subjected to shocks is to be decided by the analyst based on previous studies or dialogue with policy-makers. In effect this procedure implies that some of the parameters are being treated as "state variables".**

[2] This of course by-passes the issue of stochastic singularity altogether, since no estimation is involved.

3. **<u>Dynamic Measurement Errors :</u> This method (see Ireland (2004) for a full discussion) overcomes the stochastic singularity problem by introducing "measurement errors" in each of the observation equations in (2). These errors are presumed to follow a VAR model so that the structure (2) is modified to:**

$$f_t = \Gamma s_t + u_t \tag{5}$$

The measurement error process u_t follows an AR(1) process i.e.:

$$u_t = \mathrm{M}_1 u_{t-1} + e_t \tag{6}$$

It is assumed that (i) $E(\epsilon_t e'_t) = 0$ (ii) $e_t \sim N(0, \mathrm{V})$

Various types of assumptions have been made on the matrices M_1 and V.

- **(i) Altug (1989) and Sargent (1989) assume $\mathrm{M}_1 = 0$ while V is assumed to be diagonal. Thus the measurement errors are uncorrelated and serially uncorrelated**
- **(ii) McGrattan et al (1997) imposes a diagonal structure on both M_1 and V. Thus the measurement errors are autoregressive but uncorrelated.**
- **(iii) Finally Ireland (2004) allows the measurement errors to be correlated as well as autoregressive i.e. the matrices M_1 and V are not restricted to be diagonal. But the matrix M_1 is restricted to have all its eigenvalues within the unit circle, while V is constrained to be positive definite.**

4. **<u>Reducing the Dimensionality of the Observable vector space :</u> The approaches considered above (except calibration) overcome the "stochastic singularity" problem by increasing the number of structural shocks to match (or exceed the number of observable variables). An alternative approach would be to reduce the dimension of the space spanned by the observable vector f_t to the number of structural shocks. A naïve way to do this (which was adopted in some earlier DSGE versions) would be to restrict the set of observable variables, till equality between the number of endogenous variables and structural shocks is achieved. As pointed out by Canova et al (2014) this approach can complicate the application of full information estimation methods by leading to a highly nonlinear likelihood function, forcing the analyst to fall back upon limited information estimation methods.**

Andrews and Mikusheva (2011), Morris (2014) etc. . Andrle (2010) take an alternative approach using the DPCA (dynamic principle component analysis) of Brillinger (1981). In this approach the number of principle components considered is less than the number of shocks. The model is rotated into the subspace of the principle components space and the transformed model is estimated via Whittle's (1962) penalized likelihood method.

<u>5.Core/Non-Core Distinction :</u> From the point of view of policy applications, the approach taken by the Bank of England in developing its version of a DSGE model presents several attractive features. This model developed fully in Harrison et al (2005) and referred to as the Bank of England Quarterly Model (BEQM) distinguishes three aspects of the model viz. (i) the core model (CM) (ii) the data adjusted model (DAM) and (iii) the operational model (OM). Often, the latter two stages are referred to as the "non-core component" of the model.

The CM is a tight theoretical model solidly grounded in economic theory but does incorporate many of the institutional features and policy constraints. The DAM serves three purposes : (i) it relates the core variables to their observable counterparts (ii) it includes features such as credit market imperfections, informal sector, housing prices, agricultural sector etc. which could make the core model too complex to be tractable and (iii) it includes some relations and stylized facts for which the theoretical underpinnings are unclear (e.g. impact of monetary policy on the yield curve, factors determining the foreign exchange rate premium etc.).

The OM is the model used for actual policy purposes and incorporates extraneous information useful for policy but not amenable to formal modeling such as policymakers' judgments, consumer confidence, business surveys etc. (see Pagan (2005)). Such aspects can be modeled either by introducing specific variables (if the extraneous information can be put on a scale e.g. consumer and business confidence) or if this cannot be done (as, for example, with policymakers' or analysts' judgments) then by introducing Bayesian priors on some of the parameters in the model.

As to how the core non-core approach overcomes the stochastic singularity problem, this is shown in my recent working paper (Nachane (2016) in the context of a simple real business cycle model.

ESTIMATION OF DSGE MODELS

Basically, four estimation approaches are deployed in the DSGE context viz:

(i) Maximum likelihood

(ii) Generalized Method of Moments

(iii) Simulated Method of Moments

(iv) Indirect Inference Method

MAXIMUM LIKELIHOOD METHOD

While the logic of this procedure is straightforward, (and the maximum likelihood estimators are additionally consistent and asymptotically normal), the direct application of the method rarely succeeds in practice. Optimization in the parameter space can often fail to converge if the number of parameters is large.. The optimization hyper-surface can often be flat (and hence non-informative) about certain parameters which means that the maximization algorithm can oscillate without convergence indefinitely (see Canova and Sala (2009)).

Since a pure maximum likelihood strategy can lead to computational difficulties, a *mixed estimation* strategy is often resorted to (see De Jong et al (2000), Schorfheide (2000), Ruge-Murcia (2007) etc.). Here it is assumed that the analyst has certain prior information about a subset $\boldsymbol{\theta_e}$ of the parameter vector $\boldsymbol{\theta}$, based on economic theory, previous micro-studies, or on certain stylized empirical regularities in the data. The prior information can assume several forms but for analytical convenience, it is presumed that this information can be summarized as probability density functions referred to simply as *priors*. The set of remaining parameters about which we have no particular information can be termed as $\boldsymbol{\theta_s}$, and these are assigned "non-informative" or "diffuse" priors which are essentially flat or near flat distributions (see De Jong et al (2000)). By Bayes' theorem, it is well-known that the posterior distribution is related to the prior distributions via the following:

$$\boldsymbol{P\big((\theta|x_1 \dots x_T)\big) \propto l\big((x_1 \dots x_T|\theta)\big)P(\theta_s)} \qquad \textbf{(11)}$$

where $P(\theta_s)$ is the "diffuse" prior on θ_s, $l((x_1 \dots x_T|\theta))$ is the sample likelihood[3] conditioned on the entire vector θ , and $P((\theta|x_1 \dots x_T))$ is the posterior distribution of θ , conditioned on the observed sample. The normalizing term in the denominator of the right hand side of (11) is by Bayes theorem $P(x_1 \dots x_T)$ the unconditional joint density of the data $(x_1 \dots x_T)$ which would be unknown in most applications.

The posterior distribution is analytically intractable in most cases and has to be tackled by numerical Monte Carlo simulation.

Monte Carlo Simulation

In Monte Carlo simulation an i.i.d. sample is generated say $\{\theta^1, \theta^2, \dots \dots .. \theta^T\}$, which serves as the basis for the inference problem at hand (see Hammersley and Handscomb (1964) for an authoritative discussion). Four alternative methods are available for generating the Monte Carlo sample in the Bayesian posterior estimation problem viz:

- **(i) Acceptance/Rejection Sampling (Press et al (1992))**
- **(ii) Importance Sampling (Geweke (1989, 1997), Richard and Zhang (2007) etc.)**
- **(iii) Metropolis –Hastings Algorithm (Metropolis et al (1953), Hastings (1970), Gelfand and Smith (1990), Chib and Greenberg (1995) etc.) and**
- **(iv) Gibbs Sampling (Lange (1999), Tierney (1994) etc.) .**

---

[3] We are using $l((x_1 \dots x_T|\theta))$ to denote the likelihood, since $L((x_1 \dots x_T|\theta))$ has been used to denote the log-likelihood in (14) earlier.

Markov Chain Monte Carlo Simulation: **The induction of Markov Chains into Monte Carlo simulation following the seminal papers of Metropolis et al (1953) and Hastings (1970) gave a tremendous fillip to the entire field. Essentially, the fundamental breakthrough was the realization that under certain fairly general conditions, a Markov Chain would converge asymptotically to any distribution of interest. MCMC (Markov Chain Monte Carlo simulation) has proved extremely useful in Bayesian inference but it is more generally applicable to analyze any complex function which is analytically intractable. Basically there are two strands of MCMC which have proved most fruitful in applications viz. the Metropolis- Hastings (MH) algorithm and Gibbs Sampling. It has been the experience that these methods yield more efficient results than earlier simulation methods (such as acceptance-rejection sampling, importance sampling etc.). The basic concepts of Markov chain theory underpinning these methods can be found in standard texts such as Doob (1953), Norris (1997), Rosenblatt (1971) etc.**

Three issues are involved in these algorithms (i) to specify iterations in such a way that the associated *transition kernel* generates a *reversible* Markov chain whose *equilibrium* distribution is the *target distribution* of interest (ii) to ensure that the algorithm actually converges to the target distribution with successive draws from the candidate distribution (iii) to ensure that the convergence is fairly rapid.

Both the algorithms (M_H and Gibbs) try to address these three issues.

EVALUATION OF DSGE MODELS

Before putting any identified and estimated model to practical uses in forecasting or policy analysis, the model has to be subjected to rigorous evaluation procedures. One may distinguish three stages in the evaluation procedures of DSGE models (i) calibration (ii) sensitivity analysis and (iii) DSGE-VARs. We discuss each of these briefly below

Calibration

The earliest evaluation procedures suggested for DSGE models are based on calibration. As we have discussed above calibration imputes values to the parameters of the DSGE model based on estimates from micro-evidence, or long-run averages or from a priori theoretical considerations (as in the case of the rate of time-discounting). In early studies (e.g. Kydland and Prescott (1982), Mehra and Prescott (1985) etc.), the *population* moments $\boldsymbol{\theta}$ (of interest to the analyst) implied by these parameters and the theoretical model, were matched with the *historical moments* $\boldsymbol{\theta_T}$ thrown up by the data (T observations) . Since the sampling distribution of these historical moments are usually not known (even asymptotically), this procedure lacks statistical foundations. Gregory and Smith (1991) improve on this procedure by simulating from the DSGE model repeatedly and obtaining estimates $\boldsymbol{\theta_N}$ (N the length of the simulations). The historical moment $\boldsymbol{\theta_T}$ is used as a critical value to find the size of the test involved in comparing $\boldsymbol{\theta_N}$ with $\boldsymbol{\theta_T}$. Thus if $\boldsymbol{\theta_T}$ lies within the 95% confidence interval of $\boldsymbol{\theta_N}$, we are able to accept the DSGE model with a high degree of confidence.

Sensitivity Analysis :

The drawback of the calibration procedure is that it attributes the difference between the model results and the empirical observations to sampling errors, thus presuming that the model is true and neglecting parameter uncertainty. This deficiency is sought to be addressed by the sensitivity approaches of Kwan (1991) and Canova (1995). The latter approach is described heuristically here. Let the empirical data to be modeled be described by the vector Y_t and the corresponding model vector by the vector X_t and assume that the model is described by $X_t = f(Z_t, \beta)$ where Z_t is a vector of exogenous variables (including shocks) and β is a vector of parameters. The form of f is assumed to be known as also the probability distribution η(Z_t) of the exogenous variables. A prior $\pi(\beta)$ is imposed on the parameters and simulations are performed by drawing with replacement iid vectors β from $\pi(\beta)$ and Z_t vectors from η(Z_t).

Suppose interest attaches to the estimation of a certain moment $\mu(X_t)$ of the simulated data (which will be subsequently matched with the corresponding moment of the actual data $\mu(Y_t)$ to get an idea of the model fit). This requires knowledge of the predictive density $\Theta(X_t)$ of the simulated data. As shown by Canova (1995), this can be obtained by solving the model $X_t = f(Z_t, \beta)$ for a sufficient number of replications of Z_t *and* β (using the M-H algorithm).

If the proportion of the simulated values $\mu(X_t)$ lying within a one standard deviation band of $\mu(Y_t)$ is small or if $\mu(Y_t)$ lies in an extreme percentile of $\Theta(X_t)$, then the procedure points to model non-congruence with the data. A notable shortcoming of the Canova procedure is its asymmetric treatment of uncertainty, arising from parameter uncertainty in the model and from sampling error in the data (see De Jong et al (1996)).

DSGE-VAR:

In recent years a particular approach based on Bayesian methods is becoming increasingly popular as a method for evaluation of DSGE models. This method seems to have been introduced by De Jong et al (1996), and later developed and refined in a series of papers by Del Negro and Schorfheide (2004, 2006), and Del Negro et al (2007). It is well known in the literature that a DSGE model can be associated with a VECM (see e.g. Franchi and Juselius (2007)) in which the DSGE parameters impose cross-coefficient restrictions on the VECM parameters. Let $\boldsymbol{\theta}$ denote the parameter vector of the DSGE, $\boldsymbol{\Phi}(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}_{\boldsymbol{u}}(\boldsymbol{\theta})$ the vector of the coefficients and the var-cov matrix of the innovations of the associated reduced form VECM .

We begin with a prior distribution $\boldsymbol{p}(\boldsymbol{\theta})$ for the DSGE parameters centred at the estimated values $\boldsymbol{\theta}^*$of the DSGE model. This leads to an associated prior distribution $\boldsymbol{p}(\boldsymbol{\Phi}, \boldsymbol{\Sigma}_{\boldsymbol{u}})$ for the VECM parameters centred at $\boldsymbol{\Phi}(\boldsymbol{\theta}^*), \boldsymbol{\Sigma}_{\boldsymbol{u}}(\boldsymbol{\theta}^*)$. The precision of the prior $\boldsymbol{p}(\boldsymbol{\Phi}, \boldsymbol{\Sigma}_{\boldsymbol{u}})$ is controlled by a hyper-parameter λ ranging from λ=0 to λ=∞. The value λ=0 correspond to the case where the VECM is completely unrestricted and the value λ=∞ corresponds to the case where all the prior mass is concentrated at the values $\boldsymbol{\Phi}(\boldsymbol{\theta}^*), \boldsymbol{\Sigma}_{\boldsymbol{u}}(\boldsymbol{\theta}^*)$.

Note : The value λ=0 correspond to the case where the VECM is completely unrestricted and the value λ=∞ corresponds to the case where all the prior mass is concentrated at the values $\boldsymbol{\Phi}(\boldsymbol{\theta}^*), \boldsymbol{\Sigma_u}(\boldsymbol{\theta}^*)$.

Let Y denote the *(Txn)* matrix of T observations on n endogenous variables . we derive the conditional pdf's of $p(Y|\theta)$ and $p(Y|\lambda)$ sequentially by simulation (see Weinberg and Kyprianou (2005)). Once $p(Y|\lambda)$ is known, we can specify a grid of possible values for λ viz. $\Lambda = \{\lambda_1, \lambda_2, \ldots\ldots \lambda_q\}$ with $\lambda_1 = 0$ *and* λ_q very large (as mentioned earlier this two extreme values correspond respectively to the unrestricted VECM and the VECM with the DSGE parameter restrictions. Let $p(Y|\lambda)$ attain its maximum over the grid Λ at $\hat{\lambda}$ and consider the likelihood ratio

$$\frac{p(Y|\lambda=\hat{\lambda})}{p(Y|\lambda=\lambda_q)} \qquad \textbf{(7)}$$

A large value of $\hat{\lambda}$ and a value of the likelihood ratio (7) close to 1 is taken as evidence in favour of the DSGE model. Comparison of the impulse responses of DSGE-VECM(∞) with those of DSGE-VECM($\hat{\lambda}$) can yield useful insights into the sources of the model misspecification(see Del Negro et al (2007)).

MAIN CRITICISMS

Economic Theory Critique :

(1) *Rational Expectations.*
(2) *Representative Agent*
(3) *Complete,and Efficient Markets*
(4) *Stability of General Equilibrium*
(5) *Transversality Condition*
(6) *Ergodic Underatinty*

Econometric Critique :

(1) *Aggregation Problem*
(2) *Log-linearize and trivialize*
(3) *Inflating the "fit" by introducing ad-hoc features*

BEYOND DSGE MODELS

Two such approaches are emerging in the literature. The first is the ECONOPHYSICS literature which shifts the focus away from individual equilibria to systems equilibria and wherein evolving micro-dynamic interactions are consistent with macro equilibrium. Micro-foundations are abandoned in favour of *dimensional* analysis and the use of traditional topological methods are replaced by the methods of statistical physics (see Farmer et al (1988), Aoki and Yoshikawa (2006) and Colander (2006)).

A second, and perhaps more promising approach is the ACE (AGENT-BASED COMPUTATIONAL ECONOMICS) put forth by Epstein and Axtell (1996), Tesfatsion and Judd (2006), LeBaron and Tesfatsion (2008). ACE modeling allows for a variegated taxonomy of agents including a spectrum of cognitive features ranging from passive cognition to the most sophisticated cognitive abilities. A second important aspect of ACE modeling is that it examines the evolution of macro dynamics as the number of interacting agents increases and as their interactions become more complex. The method relies heavily on experimental designs to make inferences about the behavior of different agents. The interactions are determined by the agents' internal structures, information sets, beliefs and cognitive abilities. Agent behavior is not restrained by artificial external boundary conditions such as homogeneity, stability or transversality. Using the so-called Zipf distribution, Axtell (2001) reports a model with millions of interacting agents (see also Adamic (2011))

Nevertheless, neither of the above two approaches really validate the data in manner that could satisfy the rigorous demands of our profession. This deficiency is important and will possibly not be long in getting satisfactorily resolved. Meanwhile should we persist with the DSGE approach in spite of its problematic foundations? Solow (2010) in his testimony before the U.S. House of Representatives Committee on Science and Technology severely indicts the DSGE business " The point I am making is that the DSGE models have nothing useful to say about anti-recession policy because they have built into its essentially implausible assumptions the "conclusion" that there is nothing for macroeconomic policy to do. ….There are other traditions with better ways to do macroeconomics..". Similarly talking about the Bank of England's disillusionment with DSGE models in the aftermath of the global crisis, Buiter (2009) refers to "the chaotic re-education" at the institution.

This "re-education" could usefully incorporate three fundamental considerations viz. (i) lesser reliance on pre-selected formal models and greater scope for exploratory data analysis (ii) robustness across model specifications in policy choices and (iii) ethical responsibility of economic researchers.

One approach which is less formal (than DSGE models) and which gives greater scope for exploratory data analysis is the CVAR (co-integrated VAR) approach developed by Johansen (1996) and elaborated in Juselius (2006) and Hoover et al (2008) . It is shown in Juselius and Franchi (2007) that the assumptions underlying a DSGE model can be translated into testable hypotheses in a CVAR framework. A second approach by Del Negro and Schorfheide (2004) (DSGE-BVAR) seems even more promising. Here the estimated parameters from a DSGE

model are used as priors in an associated Bayesian VAR. A hyper-parameter λ controls the tightness with which the priors are imposed. These priors are fed into the likelihood function of the VAR to obtain the posterior distribution of the parameters. The shape of the posterior distribution for λ can help us adjudge the suitability of the tested parameters of the underlying DSGE (from the point of view of goodness-of-fit as well as model complexity). While neither of the above two approaches can claim to be perfect, they have the merit of going beyond the narrow DSGE view and allowing greater room for the data to speak.

The issue of robustness across model specifications is a largely neglected issue in the literature. In the real world policymakers are uncertain about the model(s) that they use. This uncertainty has several dimensions viz. parameter uncertainty, uncertainty about the persistence of shocks, uncertainty about the data quality etc. In such a situation what is required is a method to study the sources of model errors. The *Model Error Modeling* literature from control theory can be useful here (see Ljung (1999)). Introducing robustness considerations in economics has been studied from a different viewpoint in McCallum (1988) Hansen and Sargent (2001), Onatski and Stock (2002) etc. These ideas however have not yet filtered down to real-world policy making.

CONCLUSION

Finally, the recent global crisis has brought to the fore the ethical responsibility of the economics profession. As the financial wizards went into top gear with their innovations in the build-up to the crisis, the regulators failed to get adequate and timely warning about the potential for systemic damage in these developments, from macroeconomists in general. Are we to believe that the leading lights of our profession were simply ignorant about the dangers posed by an over-leveraged, over-securitized and skewedly-incentivized financial sector, or as is more likely they simply looked the other way ? Either view does not redound to the profession's credit. Perhaps economists should take their ethical responsibilities far more seriously than they do now and issue timely warnings to policymakers and the general public of developments which (in their opinion) are fraught with serious consequences for society at large.

Solow's (1997) characterization of academic economists as "the overeducated in search of the unknowable" is apt in the current context. Economists would be more usefully employed if instead of pursuing the Holy Grail of the true but unknown and formally perfect model, they set up a more modest agenda of studying the knowable. The lines of thinking noted briefly in the previous paragraphs (viz. the, ACE, CVAR and DSGE-BVAR models) represent precisely this line of thinking. One could not agree more with Colander ((2000), p. 131) when he sets up an agenda for those he terms the New Millenium economists as " .. search for patterns in data, try to find temporary models that fit the patterns, and study the changing nature of those patterns as institutions change".