Spanos, Aris

**Article**

# The Pre-Eminence of Theory versus the European CVAR Perspective in Macroeconometric Modeling

Economics: The Open-Access, Open-Assessment E-Journal

**Provided in Cooperation with:**
Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges

*Suggested Citation:* Spanos, Aris (2009) : The Pre-Eminence of Theory versus the European CVAR Perspective in Macroeconometric Modeling, Economics: The Open-Access, Open-Assessment E-Journal, ISSN 1864-6042, Kiel Institute for the World Economy (IfW), Kiel, Vol. 3, Iss. 2009-10, pp. 1-14,
https://doi.org/10.5018/economics-ejournal.ja.2009-10

This Version is available at:
https://hdl.handle.net/10419/27530

Economics
The Open-Access, Open-Assessment E-Journal

Vol. 3, 2009-10| April 7, 2009 | http://www.economics-ejournal.org/economics/journalarticles/2009-10

# The Pre-Eminence of Theory versus the European CVAR Perspective in Macroeconometric Modeling

*Aris Spanos*

*Virginia Tech, USA*

**Abstract**

The primary aim of the paper is to place current methodological discussions in macroeconometric modeling contrasting the 'theory first' versus the 'data first' perspectives in the context of a broader methodological framework with a view to constructively appraise them. In particular, the paper focuses on Colander's argument in his paper "Economists, Incentives, Judgement, and the European CVAR Approach to Macroeconometrics" contrasting two different perspectives in Europe and the US that are currently dominating empirical macroeconometric modeling and delves deeper into their methodological/philosophical underpinnings. It is argued that the key to establishing a constructive dialogue between them is provided by a better understanding of the role of data in modern statistical inference, and how that relates to the centuries old issue of the realisticness of economic theories.

Special issue "Using Econometrics for Assessing Economic Models"

**JEL:** B4, C1, C3

**Keywords:** Econometric methodology; 'general-to-specific'; pre-eminence of theory; cointegrated VAR; statistical adequacy; realisticness of a theory; statistical model; actual versus nominal error probabilities

**Correspondence**
*Aris Spanos, Department of Economics, Virginia Tech, 3019 Pamplin Hall, Blacksburg VA, 24061, USA; email:* ***aris@vt.edu***

*Many thanks are due to Kevin Hoover and Katarina Juselius for several valuable comments and suggestions.*

# 1 Introduction

Colander (2009) (this volume) compares and contrasts two alternative perspectives in empirical macroeconomics, and attempts to explain the extent of their influence on the discipline in terms of the incentive scheme perpetrated on the profession by US dominated journals. In broad terms his argument is that the European perspective, based primarily on the 'general-to-specific' Cointegrated Vector AutoRegressive (CVAR) approach, has been largely ignored by US dominated journals because it places observation before theory and requires researcher judgment to be part of the analysis". In contrast, the editorial boards of these journals have manifested a strong preference for the 'theory first' perspective, currently dominated by 'Dynamic Stochastic General Equilibrium' (DSGE) models, where data play only a subordinate role in 'quantifying' these models. As a result, young researchers operating in a 'publish or perish' environment would naturally avoid the European perspective because it requires hard work and judicious judgment in data modeling without any obvious professional payoff. Instead, it is rational for empirical macroeconomists to opt for the US perspective where one only needs to demonstrate technical dexterity in solving/approximating and calibrating DSGE models. Hence, the current dominance of the DSGE in empirical macro-modeling has very little to do with the superior attributes of that perspective on either substantive or empirical grounds.

Colander's incentive-based diagnosis, although broadly right-minded, does not go far enough to bring out the deeper methodological issues and the rationale underlying the two perspectives. For instance, his analysis does not explain why the US dominated journals have adopted the 'theory first' perspective in the first place, or why the European perspective places observation before theory, as he claims, knowing that such a perspective will *not* lead to publications in prestigious journals. Indeed, his 'theory first' vs. 'data first' is overly simplistic and invariably misleading because neither side will consider it as adequately characterizing their respective thesis.

The US perspective is better described as a 'Pre-Eminence of Theory' (PET) standpoint, where the data are assigned a subordinate role broadly described as 'quantifying theories presumed adequate'. In contrast, the European 'general-to-specific' CVAR perspective attempts to give data a more substantial role in the theory-data confrontation and is more accurately described as endeavoring to accomplish the goals afforded by sound practices of frequentist statistical methods in learning from data. Colander's description of the European perspective requiring 'researcher judgment' gives the misleading impression that he refers to subjective judgments and skills in statistical modeling. This is misleading because any judgement/skill/claim that can be appraised independently by other researchers is not subjective in the same sense as one's choice a prior distribution reflecting personal beliefs that nobody can question.

A crucial component of Johansen's (2007) call for assessing the premises of inference has nothing subjective about it, and the judgment/skills one needs concern the proper implementation of the Fisher-Neyman-Pearson (F-N-P) model-based statistical induction; see Cox and Hinkley (1974). In particular, he raises the question of validating the statistical premises to secure the reliability of the resulting inferences.

# 2 The Methodological Underpinnings of the Two Perspectives

## 2.1 The Pre-Eminence of Theory (PET) Perspective

Why does the pre-eminence of theory (PET) perspective currently dominate US empirical macroeconomic modeling? The short answer is that, arguably, 'it represents the *status quo*' with a long history in economics going back to Ricardo (1817). A case can be made that the PET perspective has dominated economic modeling for the last two centuries; see Spanos (2009a). The conventional wisdom underlying this perspective is that one builds simple idealized models which capture certain key aspects of the phenomenon of interest, and uses such models to gain insight concerning alternative economic policies. The role of the data is only subordinate in the sense that it can help to instantiate such models by quantifying them.

Mill (1844) articulated an early temperate form of this perspective by arguing that causal mechanisms underlying economic phenomena are too complicated – they involve too many contributing factors – to be disentangled using observational data. This is in contrast to physical phenomena whose underlying causal mechanisms are not as complicated – they involve only a few dominating factors – and the use of experimental data can help to untangle them by 'controlling' the 'disturbing' factors. Hence, economic theories can only establish general *tendencies* and not precise enough implications whose validity can be assessed using observational data. These tendencies are framed in terms of the primary causal contributing factors with the rest of the numerous (potential) disturbing factors relegated to *ceteris paribus* clauses whose appropriateness cannot, in general, be assessed using observational data. This means that empirical evidence contrary to the implications of a theory can always be explained away as due to counteracting disturbing factors. Hence, Mill (1844) rendered theory testing via observational data impossible, and attributed to the data the auxiliary role of investigating the *ceteris paribus* clauses in order to shed light on the disturbing factors which prevent the establishment of the tendencies predicted by the theory in question.

Marshall (1891) largely retained Mill's methodological stance concerning the predominance of theory over data in economic theorizing despite paying lip-service to the importance of data in economic modeling. Robbins (1935) reverted to Cairnes' (1888) more extreme version that pronounced data, more or less, irrelevant for appraising the truth of deductively established propositions. Indeed, both of them went as far as to claim that the deductive nature of economic theories bestows upon them a superior status than even physical theories because it is ultimately based on 'self-evident truths' derived by 'introspection'; according to Robbins (1935), p. 105:

"In Economics, ..., the ultimate constituents of our fundamental generalizations are known to us by immediate acquaintance. In the natural sciences they are known only inferentially. There is much less reason to doubt the counterpart in reality of the assumption of individual preferences than that of the assumption of the electron."

Robbins was well aware of the developments in statistics during the early 20th century, but dismissed their pertinence to theory appraisal in economics on the basis of the argument that such techniques are only applicable to data which can be con-

sidered as 'random samples' from a static population. Unfortunately, this argument, stemming from sheer ignorance concerning the applicability and relevance of modern statistical methods, lingers on to this day (see Mirowski, 1994). Robbins[1] was not just dismissive of any attempts to use data for theory appraisal, he jested at early attempts to quantify demand curves using an example of a 'Dr Blank investigating the demand for herrings'; see ibid., p. 107.

In modern times, echoes of that extreme version of the PET perspective can be found in Kydland and Prescott (1991):

> "The issue of how confident we are in the econometric answer is a subtle one which cannot be resolved by computing some measure of how well the model economy mimics historical data. The degree of confidence in the answer depends on the confidence that is placed in the economic theory being used." (ibid., p. 171)

Indeed, the theory being appraised should be the final arbiter:

> "The model economy which better fits the data is not the one used. Rather currently established theory dictates which one is used." (ibid., p. 174).

The great puzzle is that Kydland and Prescott never tell us how the 'currently established theory' was instituted and whether anything could ever count against it.

During the 19th and 20th centuries one can find much less extreme versions of the PET perspective where data is assigned, in principle, a less subordinate role in theory appraisal. Indeed, there is no shortage of eminent economists paying lip-service to the role of the data in economic modeling, but there is a crucial disconnect between the rhetoric and the practice; with enough perseverance one would be able to find remarks, even by the most extreme adherents to the PET standpoint, that would allude to the 'important' role of the data in economic theorizing! What was missing from economic modeling was *an appropriate modeling framework* in the context of which the theory-data confrontation can be properly applied without compromising the credibleness of either source of information. This lack of an appropriate framework is most apparent in the extensive literature initiated by Friedman (1953) concerning the realisticness of economic theories, as well as the notable methodological exchanges between Keynes and Tinbergen and Koopmans and Vinning; see Spanos (2006a).

The primary difference between the 19th and the later part of the 20th century is that the developments in statistical inference, associated with the Fisher-Neyman-Pearson (F-N-P) model-based approach that culminated in the 1930s, helped to shed illuminating light on the role of data in empirical modeling in ways which were unknown to Mill or Marshall. Unfortunately for economics, some of the key elements of the F-N-P statistical perspective, including the importance of statistical model validation, never made it into modern econometrics, primarily because the Cowles Commission literature solidified the PET perspective in econometric modeling; see Spanos (2006a).

[1] Ironically, Robbins lived long enough to regret his claims concerning "the limited predictive value of time series and suchlike statistical material":

"This part of the book, more than any other, reflects the circumstances in which it was written. It is a reaction – doubtless overdone – against the ridiculous claims of the institutionalists and the cruder econometricians." (Robbins, 1971, p. 149).

A strong case can be made (see Spanos, 2009a) that the numerous attempts to redress the balance and give data a more substantial role in theory testing were frustrated by several challenging methodological/philosophical problems bedeviling empirical modeling in economics since Ricardo (1917), the most crucial being:
(MP1) the huge gap between economic theories and the available observational data,
(MP2) the issue of assessing when a model 'accounts for the regularities in the data',
(MP3) relating statistical inferences to substantive claims, hypotheses or theories.

These same problems are currently entangling the discussion between these two perspectives rendering any dialogue between them almost impossible. Due primarily to problem (MP1), early attempts to give data a more substantive role focused on data-driven models implicitly assuming that their theoretical concepts and the available data largely coincide, and relying on goodness-of-fit measures, like the $R^2$, to assess (MP2). These attempts had disastrous consequences for empirical modeling in economics because they inadvertently contributed to the fortification of the PET perspective for a variety of reasons.

(**C1) Unreliability**. Data-driven correlation, linear regression, factor analysis and principal component analysis, relying on goodness-of-fit, have been notoriously unreliable when applied to observational data, especially in the social sciences.

(**C2) Statistical spuriousness**. The arbitrariness of goodness-of-fit measures created a strong impression that one can 'forge' significant correlations (or regression coefficients) at will, if one was prepared to persevere long enough 'mining' the data. This (mistaken) impression is almost universal among philosophers and social scientists, including economists.

(**C3) Misplaced role for substantive information**. The impression in C2 has led to widely held (but erroneous) belief that substantive subject matter (theory) information provides the only safeguard against statistical spuriousness.

Exploiting the confusions created by (C1)-(C3), the PET perspective consolidated its dominance on economic modeling and persistently charged any alternative perspective that took the data seriously, including the European CVAR approach, as yet another form of 'measurement without theory', 'data-mining' and 'hunting' for statistical significance and the like.

Admonitions and rebukes concerning the devastating effects of invoking invalid assumptions by Campos et al (2005), Johansen (2007) and Juselius and Franchi (2007) do not resonate well with the advocates of the PET perspective because they sound like a sermon they have heard many times before. To them these admonitions sound like a well-rehearsed complaint concerning the *unrealisticness* of their structural models. Indeed, numerous critics of the PET perspective have articulated the unrealisticness argument over and over again during the last two centuries, beginning with Malthus (1836) who criticized the Ricardian method as based on 'premature generalization' which occasions "an unwillingness to bring their theories to the test of experience." (ibid, p. 8).

Nevertheless, modern advocates of the PET perspective, often invoking the authority of Friedman (1953), counter that such unrealisticness is inevitable, since all models are idealizations and not faithful descriptions of reality. The abstraction/idealization argument is right-headed and perfectly legitimate at the level of the theory, but adherents of the PET perspective do not seem to appreciate the

fact that if their implicit inductive premises are invalid – vis-a-vis the data – any inferences based on such premises will be highly misleading. Indeed, in light of (C1)-(C3), the PET advocates feel that they can ignore the statistical misspecification issue and argue instead that what matters is the extent to which such models 'shed light' on the phenomenon of interest and help in formulating effective economic policies.

What they do not seem to realize is that any assessment concerning the sign, magnitude and significance of estimated coefficients, however informal, constitutes an inference whose credibility is completely undermined when the estimated model is statistically misspecified; an insight from the F-N-P model-based statistical induction.

## 2.2 The European CVAR Perspective

The European CVAR perspective has its roots in the London School of Economics (LSE) 'general-to-specific' econometrics tradition (see Sargan, 1964, Hendry, 2000), and can be best understood as an attempt to redress the balance between theory and data by avoiding both extreme positions: theory-driven vs. data-driven modeling. Having reflected on this perspective for several years, I feel that the best way to describe this European perspective is in terms of a threefold objective (aims/aspires):

(A1) to give data 'a voice of its own', independent of any economic theory,

(A2) to reliably constrain economic theorizing using the data, and

(A3) avoid 'foisting' the theory onto the data at the outset because it precludes any genuine theory testing.

In light of the huge gap between theory and data, objective (A3) renders the European CVAR perspective vulnerable to charges of 'data-mining' because any attempt to take the data seriously forces one to begin the modeling with a largely data-driven model like the Autoregressive Distributed Lag (ADL) and VAR models; see Hendry (1995). Indeed, the methodological problems (MP1)-(MP3) and the misleading impressions created by (C1)-(C3), have contributed significantly to a genuine lack of communication between the two sides, rendering any constructive dialogue between them almost impossible. For the PET advocates the European CVAR approach is another form of data-based modeling which ignores the theory, despite their declarations to the contrary, and is highly vulnerable to problems (C1)-(C3). Worse, the aims (A1)-(A3) make little sense because for them theory is the only source of legitimate information for modeling purposes.

The key to unraveling the tangled arguments separating the two perspective is provided by distinguishing between *statistical adequacy* and the *realisticness* of the structural model in question. A closer examination of the 'testing assumptions' criticism raised by the European CVAR approach (see Johansen, 2007, Juselius and Franchi, 2007), reveals that it has two separate components one of which concerns the proper application of statistical inference and the other has to do with the empirical adequacy of the structural model vis-a-vis the data in question. The first component is concerned with the validity of the probabilistic assumptions comprising the inductive premises for inference. It's only the second component that relates to the centuries old *realisticness* criticism (see Maki, 2000). Hence, the advocates of

the PET perspective cannot deflect or sidestep the statistical inadequacy criticism by invoking their arguments against the realisticness of a theory criticism; the two issues are fundamentally different.

For a proper understanding of these two components and their respective roles one needs a methodological framework where these and related issues are clearly brought out. A framework that can be used to elucidate the strengths and weaknesses of both perspectives and provide the basis for a constructive dialogue between them. The same framework should also offer suggestions on how one might be able to address the methodological problems (C1)-(C3) mentioned above, as well as accommodate the threefold objective (A1)-(A3) of the European perspective.

# 3 An All-Encompassing Methodological Framework

Spanos (1986), p. 17, proposed an all-encompassing methodological framework (Figure 1), devised to enable the modeler to bridge the gap between theory and data using a sequence of interconnected models with a view to delineate and probe for the potential errors at different stages of modeling; see Mayo (1996) for a similar proposal.

The key to unraveling the testing of assumptions argument is provided by drawing a clear distinction between *substantive* and *statistical assumptions* because their respective validity has very different implications for inference. The substantive assumptions pertain to the realisticness issue, but the statistical assumptions pertain to the (statistical) reliability of inference. This is because when any of the statistical assumptions are invalid for data $\mathbf{Z}_0$, inferences based on the estimated model are often unreliable because the *nominal* and *actual error probabilities* are likely to be different. The surest way to lead an inference astray is to apply a .05 significance test when the actual type I error is closer to 1.0; see Spanos and McGuirk (2001).

The crucial problem in econometric modeling is that foisting the substantive information on the data by estimating the structural model $\mathcal{M}_{\boldsymbol{\varphi}}(\mathbf{z})$ directly, is invariably an injudicious strategy because statistical specification errors are likely to undermine the prospect of reliably evaluating the relevant errors for primary inferences. When modeling with observational data, the estimated $\mathcal{M}_{\widehat{\boldsymbol{\varphi}}}(\mathbf{z})$ is often both statistically and substantively inadequate, and one has no way to delineate the two; is the theory wrong or are the (implicit) inductive premises invalid for data $\mathbf{Z}_0$? To avert this impenetrable quandary, the modeling framework in Figure 1 distinguishes, *ab initio*, between statistical and substantive information and then allows for bridging the gap between them by a sequence of interconnecting models which enable one to delineate and probe for the potential errors at different stages of modeling. From the theory side, the substantive information is initially encapsulated by a theory model and then modified into a structural one $\mathcal{M}_{\boldsymbol{\varphi}}(\mathbf{z})$ to render it estimable with data $\mathbf{z}_0$. From the data side, the statistical information is distilled by a statistical

model $\mathcal{M}_{\boldsymbol{\theta}}(\mathbf{z})$ whose parameterization is chosen with a view to render $\mathcal{M}_{\boldsymbol{\varphi}}(\mathbf{z})$ a reparametrization/restriction thereof.

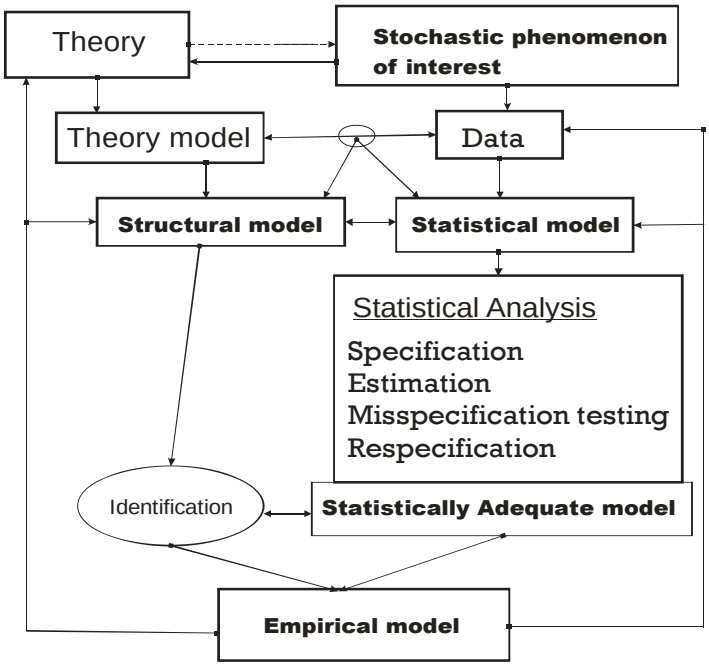


Figure 1: An Empirical Modeling Framework

Distinguishing between substantive and statistical assumptions is not as straight forward as it might seem at first sight. The problem can be seen in Ireland (2004) where the assumptions invoked: (1) all structural parameters are constant over time, (2) total factor productivity is driving the system, (4) log output, consumption, and capital are trend-stationary, (5) labor is stationary, (6) labor augmented technological progress follows a linear trend which influences the other variables identically, (7) the observable variables follow a VAR(1) process, (8) the errors are NIID, constitute a mixture of substantive and statistical assumptions; see Juselius and Franchi (2007).

The initial separation depends on having a clear-cut distinction between a structural $\mathcal{M}_{\boldsymbol{\varphi}}(\mathbf{z})$ and a statistical model $\mathcal{M}_{\boldsymbol{\theta}}(\mathbf{z})$ where the former is viewed as an *estimable* form of a theory model (hence, built on substantive information) in view of the available data $\mathbf{Z}_0$, and the latter as a purely probabilistic construal whose structure depends solely on the statistical information contained in the data $\mathbf{Z}_0{:=}(\mathbf{z}_t,\ t{=}1,2,...,n)$; see Spanos (1986). The latter is accomplished by viewing the statistical model as a particular parameterization of a generic vector stochastic process $\{\mathbf{Z}_t,\ t{\in}\mathbb{N}\}$ whose probabilistic structure is chosen so as to render data $\mathbf{Z}_0$ a 'truly typical realization' of this process. The particular parameterization of $\{\mathbf{Z}_t,\ t{\in}\mathbb{N}\}$ is selected so as to enable one to embed the structural model in its context.

**Table 1- Normal Vector Autoregressive (VAR(1)) Model**

| | | |
|---|---|---|
| Statistical GM: | | $\mathbf{Z}_t = \mathbf{a}_0 + \mathbf{A}_1^\top \mathbf{Z}_{t-1} + \mathbf{u}_t,\ t\in\mathbb{N},$ |
| [1] | Normality: | $D(\mathbf{Z}_t\vert\ \mathbf{Z}_{t-1}^0;\boldsymbol{\theta})$, for $\mathbf{Z}_{t-1}^0{:=}(\mathbf{Z}_{t-1},...\mathbf{Z}_1)$, |
| [2] | Linearity: | $E(\mathbf{Z}_t\vert\sigma(\mathbf{Z}_{t-1}^0))= \mathbf{a}_0 + \mathbf{A}_1^\top \mathbf{Z}_{t-1}$ , |
| [3] | Homoskedasticity: | $Var(\mathbf{Z}_t\vert\sigma(\mathbf{Z}_{t-1}^0)) = \Omega$ is free of $\mathbf{Z}_{t-1}^0$, |
| [4] | Markov dependence: | $\{\mathbf{Z}_t,\ t\in\mathbb{N}\}$ , is a Markov process |
| [5] | t-invariance: | $\boldsymbol{\theta}{:=}(\mathbf{a}_0, \mathbf{A}_1, \Omega)$ are t-invariant, |
| | $\mathbf{a}_0 = \boldsymbol{\mu} - \mathbf{A}_1^\top \boldsymbol{\mu},\ \ \mathbf{A}_1 = \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1,\ \ \Omega := \boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1^\top \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1$ | |

**Example**. In the case where the process $\{\mathbf{Z}_t,\ t\in\mathbb{N}\}$ is Normal, Markov and Stationary one can show that it can be parameterized in the form of the VAR(1) model, as specified in Table 1; see Spanos (1995).

However, depending on the structural model in question, one could choose another parameterization of the same process represented by the *Dynamic Linear Regression* model whose statistical GM, for $\mathbf{Z}_t{:=}(\mathbf{X}_t, \mathbf{y}_t)$, takes the form:

$$\mathbf{y}_t = \boldsymbol{\beta}_0 + \mathbf{B}_0^\top \mathbf{y}_{t-1} + \mathbf{B}_1^\top \mathbf{x}_t + \mathbf{B}_2^\top \mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t,\ t\in\mathbb{N},$$

and its parameters $\boldsymbol{\psi}{:=}(\boldsymbol{\beta}_0, \mathbf{B}_0, \mathbf{B}_1, \mathbf{B}_2, \mathbf{V})$ constitute re-parameterization of $\boldsymbol{\theta}{:=}(\mathbf{a}_0, \mathbf{A}_1, \Omega)$ in the sense that $\boldsymbol{\psi}{=}\mathbf{H}(\boldsymbol{\theta})$; see Spanos (1986).

It turns out that the sequence of models, theory, structural (estimable) and statistical, provides a way to foreground as well as address problem (MP1) raised above. The separation is particularly crucial because statistical adequacy [the validity of the statistical assumptions vis-a-vis data $\mathbf{Z}_0$] is a sufficient condition for the reliability of inference. Indeed, one cannot even pose questions of substantive adequacy [does the structural model capture the key features of the phenomenon of interest?] unless statistical adequacy has been secured first. This is because statistical adequacy ensures that the relevant error probabilities are ascertainable since the actual approximate closely the nominal ones; see Spanos (2006a).

The notion of statistical adequacy replaces goodness-of-fit as *the* criterion for assessing whether a fitted model 'accounts for the regularities in the data', addressing problem (MP2), and at the same time shedding ample light on the problems (C1)-(C3) misleadingly invoked by the PET advocates; see Spanos (2009a). Statistical adequacy is achieved by applying thorough *misspecification testing* to probe effectively the different ways the model assumptions (e.g. [1]-[5] in Table 1) might be misspecified; see Spanos (2000). Although the effectiveness of misspecification testing requires judicious use of graphical techniques, there is nothing subjective about the judgment needed to validate a statistical model; see Mayo and Spanos (2004).

The crucial issue here is that statistical adequacy is separate from any issues pertaining to the realisticness or the substantive adequacy of the structural model in question. In particular, statistical misspecification cannot be fended off using locutions like: "All models are misspecified, to a greater or lesser extent, because they are by necessity mere approximations, and slight departures from assumptions will only lead to minor deviations from the optimal inferences." Such locutions are

highly misleading because even seemingly minor misspecifications can yield major discrepancies between actual and nominal error probabilities; Spanos (2005).

A statistically adequate model $\mathcal{M}_{\boldsymbol{\theta}}(\mathbf{z})$ provides a sound basis for appraising the relevant structural model $\mathcal{M}_{\boldsymbol{\varphi}}(\mathbf{z})$, where the two are related via an implicit function $\mathbf{G}(\boldsymbol{\varphi},\boldsymbol{\theta})=\mathbf{0}$, where $\boldsymbol{\varphi}\in\Phi$, and $\boldsymbol{\theta}\in\Theta$, denote the structural and statistical parameters, respectively. This provides a link between $\mathcal{M}_{\boldsymbol{\theta}}(\mathbf{z})$ and the phenomenon of interest via $\mathcal{M}_{\boldsymbol{\varphi}}(\mathbf{z})$, invariably known as *identification*: does $\mathbf{G}(\boldsymbol{\varphi},\boldsymbol{\theta})=\mathbf{0}$ define $\boldsymbol{\varphi}$ *uniquely* in terms of $\boldsymbol{\theta}$? Often, there are more statistical than structural parameters, and that enables one to test the *overidentifying restrictions*:

$$H_0:\mathbf{G}(\boldsymbol{\varphi},\boldsymbol{\theta})=\mathbf{0},\text{ vs. }H_1:\mathbf{G}(\boldsymbol{\varphi},\boldsymbol{\theta})\neq\mathbf{0}. \tag{1}$$

Rejection of the null provides evidence against the empirical adequacy of the structural model vis-a-vis data $\mathbf{Z}_0$. This view of identification differs from the traditional textbook notion (see Kennedy, 2008) in so far as it requires that the underlying $\mathcal{M}_{\boldsymbol{\theta}}(\mathbf{z})$ (the reduced form) be validated vis-a-vis data $\mathbf{Z}_0$ to secure the thustworthiness of the link between $\mathcal{M}_{\boldsymbol{\varphi}}(\mathbf{z})$ and the phenomenon of interest; Spanos (1990).

Appraising the overidentifying restrictions in (1) requires one to go beyond the statistical significance to assess the substantive significance in order to adequately address problem (MP3) above by circumventing the fallacies of acceptance and rejection. This comes in the form of a post-data evaluation of inference to determine the discrepancy from the null warranted by data $\mathbf{Z}_0$ using severe testing reasoning; see Mayo and Spanos (2006), Spanos (2006b). Indeed, the modeling framework in Figure 1 can be used to address all three methodological/philosophical problems (MP1)-(MP3).

Viewed in the context of Figure 1, the PET perspective often ignores the right hand side; the statistical analysis steps leading to a statistically adequate model. Quantifying the structural model $\mathcal{M}_{\boldsymbol{\varphi}}(\mathbf{z})$ directly usually results in an estimated (or calibrated) model which is both statistically and substantively inadequate, but without any way to separate or eliminate the different sources of error arising at the different stages of modeling; theory, structural and statistical models. Hence, any inference based on such quantified structural models will be invariably misleading. As a methodology of learning from data, it does not live up to standards of scientific objectivity that requires its theories be thoroughly tested against data; see Hoover (2006), Spanos (2009a).

A crucial consequence of distinguishing between statistical and substantive information, *ab initio*, is that the framework in Figure 1 encourages the empirical discovery process. One does not need to have a full-blown structural model like the DSGE to begin the empirical modeling process, as the Cowles Commission approach would have us believe. One can begin with low level theories (however vague) that identify certain potentially relevant variables $\mathbf{Z}_t$, and then use a statistically adequate model $\mathcal{M}_{\boldsymbol{\theta}}(\mathbf{z})$ to reliably constrain economic theorizing with a view to develop more adequate structural models for the phenomenon of interest. Without underestimating the difficulties associated with the empirical discovery process, this creates the common ground for reconciling the two perspectives.

The European CVAR perspective arguably ignores the left hand side of Figure 1 by relying on some low level theory to begin the modeling process. Once the

data $\mathbf{Z}_0$ have been chosen on the basis of a theory or theories, one can proceed to specify a statistical model, like a VAR (Table 1), in terms of the probabilistic structure of the underlying stochastic process $\{\mathbf{Z}_t,\ t\in\mathbb{N}\}$. This enables one to carry out the statistical analysis without any references to the structural model until a statistically adequate model is reached. At that stage one can proceed to impose *data-induced restrictions*, like the ones implied by *cointegration*, and attempt to relate the restricted model to certain low-level theories associated with the long-run steady-state and/or equilibrium-correction states; see Hendry (1995), Johansen (1996), Juselius (2006).

This leaves the European CVAR perspective vulnerable to the charge that their use of substantive information is rather superficial because the data-induced restrictions are only tangentially connected to economic theory. In their defense, advocates of the European perspective are likely to offer a plethora of evidence that the PET strategy give rise to structural models, like Ireland's (2004) DSGE model, which are invariably empirically incongruous; see Juselius and Franchi (2007), Hoover et al (2008).

# 4 Can the Two Perspectives Be Reconciled?

Viewing both perspectives in the context of the modeling framework in Figure 1, the advocates of the European CVAR perspective need to go the extra mile to bridge the gap between theory and data by developing structural models beyond the ones associated with data-induced restrictions. On the other hand, the adherents to the PET perspective need to develop structural models that account for the statistical regularities in the data. Statistically adequate models can be used to give data a voice of its own, to reliably constrain economic theorizing, and, one hopes, help direct the search toward more adequate structural models.

Taking Ireland's (2004) DSGE model as an example, one needs to derive explicitly the implicit reduced form and state its probabilistic assumptions (analogous to assumptions [1]-[5] in Table 1) by viewing it as a statistical model; a parameterization of the probabilistic structure of the process $\{\mathbf{Z}_t,\ t\in\mathbb{N}\}$ underlying data $\mathbf{Z}_0$. Thorough misspecification testing will determine if the latter is statistically adequate or not. Based on past experience, it is highly unlikely that such a model will turn out to be statistically adequate; see Juselius and Franchi (2007), Hoover et al (2008). This, by itself, provides empirical evidence against the structural model as it stands, and a respecification aiming to account for the statistical regularities in data $\mathbf{Z}_0$ is called for.

It is important to stress that respecification in this context does not refer to 'error-fixing' widely used in traditional textbook econometrics, but postulating more appropriate probabilistic structure for $\{\mathbf{Z}_t,\ t\in\mathbb{N}\}$ that would render data $\mathbf{Z}_0$ a typical realization thereof. This is because the traditional 'error-fixing' strategies, such as error-autocorrelation correction and heteroskedasticity/autocorrelation consistent standard errors (see Kennedy, 2008), often render statistical unreliability worse, not better; see Spanos and McGuirk (2001), Spanos (2006a).

Assuming one can find such a respecified statistical model, it can provide the basis for improving the original structural model using modifications that take into

account the statistical regularities as described by the statistically adequate model. In a sense, the latter demarcates 'what there is to be explained' by potential structural models that aspire to be empirically adequate. This process might require several iterations before such a model is reached.

# 5 Conclusion

Real progress in learning from data about economic phenomena of interest can be expected when economic modelers face squarely the formidable difficulties in addressing all three methodological problems (MP1)-(MP3) mentioned above. The main message from the above discussion is that these challenging problems can be addressed in the context of the modeling framework shown schematically in Figure 1. The key is provided by recognizing that, although both substantive and statistical information play crucial roles in learning from data, their respective roles in empirical modeling need to be delineated and properly reconciled. The proposed reconciliation is achieved in the broader context of bridging the gap between theory and data using a sequence of interconnecting models (Figure 1). This framework creates common ground for a constructive dialogue between economic theorists and econometricians that could give rise to 'learning from data' about economic phenomena of interest.

What are the prospects that such a constructive dialogue will begin any time soon? Despite the gloomy picture painted above, I remain optimistic that the new generation of econometricians will eventually grow out of esteeming technical dexterity and begin to reflect on the serious methodological issues undermining the trustworthiness of the evidence produced by the prevailing econometric modeling practice; see Spanos (2009b). The primary motive for this change is likely to be that, as things stand, the prospect of econometric modeling losing its credibility as a serious scientific field vis-a-vis other scientists as well as policy makers looms large; see Spanos (2008).

# References

Cairnes, J. E. (1888), *The Character and Logical Method of Political Economy*, Reprints of Economic Classics, 1965, Augustus, M. Kelley, NY.

Campos, J., N. R. Ericsson, and D. F. Hendry (2005) (eds.), *General-to-Specific Modeling*, Vol. 1-2, Edward Elgar, Northampton, MA.

Colander, D. (2009), "Economists, Incentives, Judgement, and the European CVAR Approach to Macroeconometrics," *Economics—The Open-Access, Open-Assessment E-Journal*,Vol. 3, 2009-9, http://www.economics-ejournal.org/economics/journalarticles/2009-9.

Cox, D. R. and D. V. Hinkley (1974), *Theoretical Statistics*, Chapman & Hall, London.

Friedman, M. (1953), "The Methodology of Positive Economics," pp. 3-43 in *Essays in Positive Economics*, Chicago University Press, Chicago.

Hendry, D. F. (1995), *Dynamic Econometrics*, Oxford University Press, Oxford.

Hendry, D. F. (2000), *Econometrics: Alchemy or Science?* 2nd ed., Blackwell, Oxford.

Hoover, K. D. (2006), "The Methodology of Econometrics," pp. 61-87, in *New Palgrave Handbook of Econometrics*, vol. 1, ed. T. C. Mills and K. Patterson, Macmillan, London.

Hoover, K. D., S. Johansen and K. Juselious (2008), "Allowing the Data to Speak Freely: The Macroeconometrics of the Cointegrated Vector Autoregression," *American Economic Review*, 98: 251-255.

Ireland, P. 2004. "A Method for taking Models to the Data" *Journal of Economic Dynamics and Control,* 28: 1205-26.

Johansen, S. (1996), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.

Johansen, S. (2007), "Confronting the Economic Model with the Data" in D. Colander, *Post Walrasian Macroeconomics: Beyond the DSGE Model*. Cambridge University Press.

Juselius, K. (2006), *The Cointegrated VAR Model: Methodology and Applications*, Oxford University Press, Oxford.

Juselius, K. and Franchi, M. (2007), "Taking a DSGE Model to the Data Meaningfully." *Economics–The Open-Access, Open-Assessment E-Journal*, No. 2007-4, http://www.economics-ejournal.org/ economics/journalarticles/2007-4.

Kennedy, P. (2008), *A Guide to Econometrics*, 6th edition, MIT Press, MA.

Kydland, F. and P. Prescott (1991), "The Econometrics of the General Equilibrium Approach to Business Cycles," *Scandinavian Journal of Economics*, 93: 161-178.

Maki, U. (2000), "Reclaiming Relevant Realism," *Journal of Economic Methodology*, 7: 109-125.

Malthus, T. R. (1936), *Principles of Political Economy*, Reprints of Economic Classics, 1986, Augustus, M. Kelley, NJ.

Marshall, A. (1891/1920), *Principles of Economics*, 8th ed., McMillan, London.

Mayo, D. G. (1996), *Error and the Growth of Experimental Knowledge*, The University of Chicago Press, Chicago.

Mayo, D. G. and Spanos, A. (2004), "Methodology in Practice: Statistical Misspecification Testing," *Philosophy of Science*, 71: 1007–25.

Mayo, D. G. and Spanos, A. (2006), "Severe testing as a basic concept in a Neyman–Pearson philosophy of induction," *British Journal for the Philosophy of Science*, 57: 323–57.

Mill, J. S. (1844), *Essays on some Unsettled Questions of Political Economy*, Reprints of Economic Classics, 1974, Augustus, M. Kelley, NJ.

Mirowski, P. (1994), "What are the questions?" pp. 50- 74 in Backhouse, R. E. (ed) (1994), *New Directions in Economic Methodology*, Routeledge, London.

Ricardo, D. (1817), *Principles of Political Economy and Taxation, vol. 1 of The Collected Works of David Ricardo*, ed. P. Sraffa and M. Dobb, Cambridge University Press, Cambridge.

Robbins, L. (1935), *An Essay on the Nature and Significance of Economic Science*, 2nd ed., McMillan, London.

Robbins, L. (1971), *Autobiography of an Economist*, McMillan, London.

Sargan, J. D., (1964), "Wages and prices in the United Kingdom: a study in econometric modelling," pp 25–63 in Hart, P. E., Mills, G. and Whitaker, J. K., *Econometric Analysis for National Economic Planning*, Butterworths, London.

Spanos, A. (1986), *Statistical Foundations of Econometric Modelling*, Cambridge University Press, Cambridge.

Spanos, A. (1990), "The Simultaneous Equations Model Revisited: Statistical Adequacy and Identification," *Journal of Econometrics*, 44: 87–108.

Spanos, A. (1995), "On theory testing in econometrics: modeling with nonexperimental data," *Journal of Econometrics*, 67: 189–226.

Spanos, A. (2000), "Revisiting Data Mining: 'hunting' with or without a license," *The Journal of Economic Methodology*, **7**: 231-264.

Spanos, A. (2005), "Misspecification, Robustness and the Reliability of Inference: the simple t-test in the presence of Markov dependence," Working Paper, Virginia Tech.

Spanos, A. (2006a), "Econometrics in retrospect and prospect," In *New Palgrave Handbook of Econometrics*, vol. 1, ed. T. C. Mills and K. Patterson, Macmillan, London.

Spanos, A. (2006b), "Revisiting the omitted variables argument: substantive vs. statistical adequacy," *Journal of Economic Methodology*, 13: 179–218.

Spanos, A. (2008), "Statistics and Economics," pp. 1129-1162 in the *New Palgrave Dictionary of Economics*, 2nd edition, Eds. S. N. Durlauf and L. E. Blume. Palgrave Macmillan, London.

Spanos, A. (2009a), "Theory Testing in Economics and the Error Statistical Perspective," forthcoming in *Error and Inference*, D. G. Mayo and A. Spanos (2009b), Cambridge University Press, Cambridge.

Spanos, A. (2009b), "Philosophy of Econometrics," forthcoming in the *Handbook of the Philosophy of Science*, edited by D. Gabbay, P. Thagard, and J. Woods, Elsevier, North Holland.

Spanos, A. and McGuirk, A. (2001), "The model specification problem from a probabilistic reduction perspective," *Journal of the American Agricultural Association*, 83: 1168–76.