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

Discussion Paper 2008-25
June 24, 2008

The'Pre-Eminence of Theory' versus the 'General-to-Specific' Cointegrated VAR Perspectives in Macro-Econometric Modeling

Aris Spanos

Department of Economics, Virginia Tech

Please cite the corresponding journal article:
http://www.economics-ejournal.org/economics/journalarticles/2009-10

Abstract:

The primary aim of the paper is to place current methodological discussions on empirical 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 Empirical Work" relating to the two different perspectives in Europe and the US that are currently dominating empirical macro-econometric modeling and delves deeper into their methodological/philosophical foundations. 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.

JEL: B4, C1, C3

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

Correspondence:
Aris Spanos, Department of Economics, Virginia Tech, Blacksburg VA, 24061, USA; e-mail: aris@vt.edu
Thanks are due to Bent Nielsen for several suggestions which helped to improve the discussion in the paper.

1 Introduction

Colander (2008) in a paper entitled "Economists, Incentives, Judgment and Empirical Work", makes a reasonably convincing case that empirical macro-economics in Europe and the United States (US) are currently dominated by two alternative perspectives, he calls the 'general-to-specific' cointegrated VAR approach and the 'theory comes first', respectively, and attempts to explain the differences primarily in terms of the incentive scheme perpetrated on the profession by US dominated journals.

In broad terms Colander's (2008) argument is that the European perspective is largely incompatible with the incentives facing a researcher whose primary objective is to publish the maximum number of papers that can make it through the refereeing process in US dominated journals. This is because such journals have a strong preference for the 'theory comes first' perspective where the empirical modeling is decidedly theory-driven and data play only an auxiliary role in helping to quantify structural models. In contrast, "the European 'general-to-specific' cointegrated VAR approach places observation before theory and requires researcher judgment to be part of the analysis." (ibid. p. 1).

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. Indeed, adopting the European perspective could lead to their academic demise since the only currency for academic promotion and success is increasingly the number of papers published in such journals. This also increasingly applies to established European econometricians whose careers depend more and more on publishing in US dominated journals because departments in Europe are actively competing in world-wide rankings which are also towered by publications in such journals. Hence, it is rational for empirical macroeconomists and econometricians in both Europe and the US to opt for the easier to implement US perspective where one needs to demonstrate technical dexterity in solving 'Dynamic Stochastic General Equilibrium' (DSGE) models, regardless of their empirical relevance, with only superficial knowledge of current econometric techniques at the level of being able to use a computer package to 'quantify' such models.

Colander's diagnosis, as far as it goes, is broadly right-minded, but does not go far enough in bringing out the deeper methodological/philosophical issues that underlay the two perspectives, and does not explain the rationale for their espousal. For instance, his analysis does not explain why the US dominated journals have adopted the 'theory comes first' perspective in the first place, or why pioneers of the European perspective place observation before theory, knowing that such a perspective will *not* lead to publications in 'prestigious' journals. Indeed, his 'theory comes first' vs. 'observation comes first' is overly simplistic and somewhat misleading because neither side will consider it as characterizing their respective thesis adequately. Moreover, his description of the European perspective requiring 'researcher judgment' gives the (potentially) misleading impression that he refers to *subjective* judgments and non-transferable *personal skills* in analyzing data.

The US stance can be more aptly described as a 'pre-eminence of theory' perspective where the empirical modeling is driven by a theory and the data are given a subordinate role broadly described as 'quantifying theories presumed adequate'. In contrast, the European 'general-to-specific' cointegrated VAR perspective attempts to give data a more substantial role in the theory-data confrontation, and is better described as 'endeavoring to accomplish the goals afforded by sound practices of frequentist statistical methods in learning from data'; see Spanos (2008a). As argued below, a crucial component of the call for assessing the premises of inference by Johansen (2006) and Juselius and Franchi (2007) has nothing subjective about it, because the judgment one needs for such an assessment relies primarily on a sound understanding of the Fisher-Neyman-Pearson approach to inference; see Cox and Hinkley (1974).

2 The two perspectives and their methodological foundations

2.1 The pre-eminence of theory perspective

Why does the 'pre-eminence of theory' perspective currently dominates US empirical macro-modeling? The simple answer is that it represents the *status quo* with a very long history in economics going back to Ricardo (1817). A case can be made that it has been the dominating modeling perspective in economics for almost two centuries; see Spanos (2008a). The conventional wisdom underlaying this perspective is that one builds simple models (by necessity unrealistic) which capture certain key aspects of the phenomenon of interest, and uses them to gain insight concerning alternative economic policies. The role of the data is only auxiliary 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 certain physical phenomena whose underlying causal mechanisms 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 (potentially) numerous disturbing factors relegated to *ceteris paribus* clauses whose appropriateness cannot 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 (1944) 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 a particular theory.

A more extreme version of the pre-eminence of theory was articulated by Cairnes

(1875/1888) whose reasoning was that, given the 'self-evident truth' of the initial postulates of political economy (since they are based on introspection), and the deductive validity of the propositions which follow, the question of verification using data does not even arise because *the truth of the premises ensures the truth of the conclusions*! As a result, the deductive nature of economic theories bestows upon them a superior status that even physical theories. This is because the premises of Newtonian mechanics are not 'self-evident truths', as in economics, but mere inductive generalizations that need to rely on experimentation and inductive inferences which are known to be fallible; see ibid., pp. 72-94.

Later Marshall (1891/1920) largely retained Mill's methodological stance concerning the pre-eminence of theory over data in economic theorizing, but Robbins (1935), reverted to Cairnes' more extreme version, and went as far as to pronounce data, more or less, irrelevant in appraising the truth of deductively established propositions (ibid., p. 105).

In modern times, echoes of this extreme version of the pre-eminence of theory 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).

To what extent can one explain, or at least rationalize, the longevity of the pre-eminence of theory perspective in reigning empirical modeling in economics? 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 frequentist approach, helped to shed crucial light on the role of data in empirical modeling in ways which were unknown to Ricardo, Mill, Marshall or even Robbins. Unfortunately for economics, some of the key elements of the Fisher-Neyman-Pearson perspective never made it into modern econometrics, and as a result, the abuse of these inferential tools is widespread in current empirical modeling in economics. The extent of the problem is severe enough to render the overwhelming majority of published empirical results untrustworthy; see Spanos (2006a).

Early attempts to redress the balance and give data a more substantial role in theory testing were frustrated by several inveterate methodological/philosophical problems, including:

(i) the huge gap between economic theories and the available observational data,
(ii) the difficulty of assessing when a fitted model 'accounts for the regularities in the data', and
(iii) relating statistical inference results to substantive claims or theories.

These problems have also played an important role in rendering any constructive dialogue between the above two perspectives very difficult. Due primarily to problem (i) early attempts to give data a more substantial role focused by necessity on data-driven models and relied (misguidedly) on goodness-of-fit measures, like the R^2 and likelihood-based metrics, to address (ii). These attempts had disastrous consequences for empirical modeling in economics because they inadvertently contributed to the fortification of the 'pre-eminence of theory' perspective giving rise to several highly misleading impressions that dominate current thinking in economics.

(**I) 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.

(**II) 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.

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

As a result of (I)-(III), advocates of the prevailing pre-eminence of theory perspective in economics have persistently charged any form of modeling that takes the probabilistic structure of the data seriously with 'measurement without theory', 'data-mining' and 'hunting' for statistical significance. Indeed, any approach that does not adhere to the 'pre-eminence of theory' perspective, including the 'general-to-specific' cointegrated VAR approach, is suspiciously viewed as yet another form of 'data-mining', and thus dismissed because the conventional wisdom is that, in light of (I)-(III), one 'can torture any data set sufficiently to admit to any claim or assertion'.

Admonitions and rebukes concerning the devastating effects of relying on invalid assumptions by Campos, Ericsson and Hendry (2005), Johansen (2006) and Juselius and Franchi (2007) do not resonate well with the advocates of the pre-eminence of theory perspective. To them these admonitions sound like a well-rehearsed complaint concerning the *unrealisticness* of their structural models which they have heard many times before. Indeed, numerous critics of the pre-eminence of theory 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). However, the advocates of the pre-eminence of theory feel that such unrealisticness is inevitable, since all models are unrealistic by necessity; invoking the authority of Friedman (1953). Unlike Friedman, they will not use good predictive ability as a primary criterion for an empirically adequate model. Instead, they are likely to argue that what matters is not whether the model is realistic enough, but the extent to which such a theory-driven model 'sheds light' on the phenomenon of interest and helps in formulating effective economic policies. Of course, in their eyes the proposed structural model constitutes the epitome of illumination and provides clear support for their preferred policies, that's why it deserves publication in a prestigious journal!

But, what about the data? For a hard-core advocate of the pre-eminence of theory perspective a structural model is considered *empirically adequate* as long as it "is not obscenely at variance with the data" (Sargent, 1976, p. 233).

2.2 The 'general-to-specific' cointegrated VAR perspective

The European 'general-to-specific' cointegrated VAR perspective has its roots in the LSE tradition (see Sargan, 1964, Hendry, 2000, 2003), and can be best understood as an attempt to redress the balance between theory and data by avoiding both extreme positions: theory-driven and data-driven modeling. Having reflected on this perspective for several years, I feel the best way to describe this European perspective is in terms of its threefold objective:

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

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

(c) to find ways to relate statistical models to theories, without 'forcing' the theory onto the data preempting any genuine theory testing.

In light of the chronic problem (i), objective (c) renders the 'general-to-specific' perspective vulnerable to charges of 'data-mining' because unless one finds ways to bridge the theory-data gap, the only open venue when one wants to take the data seriously is to begin with what appears to a pre-eminence of theory advocate as a data-driven model like a VAR. The reality is that, in light of the inveterate problems (i)-(iii) and the misleading impressions created by (I)-(III), there is a genuine lack of communication between the two sides primarily due to the difficulty in dealing with problem (i), rendering any constructive dialogue between the two sides almost impossible; each side feeling that they are facing hearing-impaired critics.

For the pre-eminence of theory advocates the 'general-to-specific' cointegrated VAR approach is another form of data-based modeling which ignores the theory, despite their proclamations to the contrary, and is highly susceptible to problems (I)-(III). Indeed, for such advocates the objectives (a)-(c) make little sense because theory is the only source of legitimate information relevant for understanding and modeling economic phenomena of interest. Indeed, some advocates go as far as to argue that what distinguishes econometrics from statistics is that the former is theory-driven and rightly so. Moreover, they do not understand why critics do not recognize that structural models have to be, by necessity, unrealistic.

The 'testing assumptions' criticism raised by the 'general-to-specific' cointegrated VAR approach, however, is somewhat different (see Johansen, 2006) from the centuries old realisticness criticism, and cannot be deflected or sidestepped as easily. This criticism has two separate components one of which concerns the proper application of modern frequentist statistical inference and the other has to do with the empirical adequacy of the structural model in question; the former has nothing to do with the realisticness issue as such.

For a proper understanding of these two components and their respective roles one needs a methodological framework where such issues are clearly brought out and can accommodate both perspectives in a way that highlights the chronic problems (i)-(iii) mentioned above, as well as suggest ways they might be potentially addressed.

3 An all-encompassing methodological framework

A methodological framework that can be used to analyze the two perspectives discussed above was proposed in Spanos (1986), shown schematically in fig. 1; see ibid., p. 21. This framework can provide the basis for establishing a creative dialogue between the two perspectives and, hopefully, give rise to real progress in empirical modeling in economics.

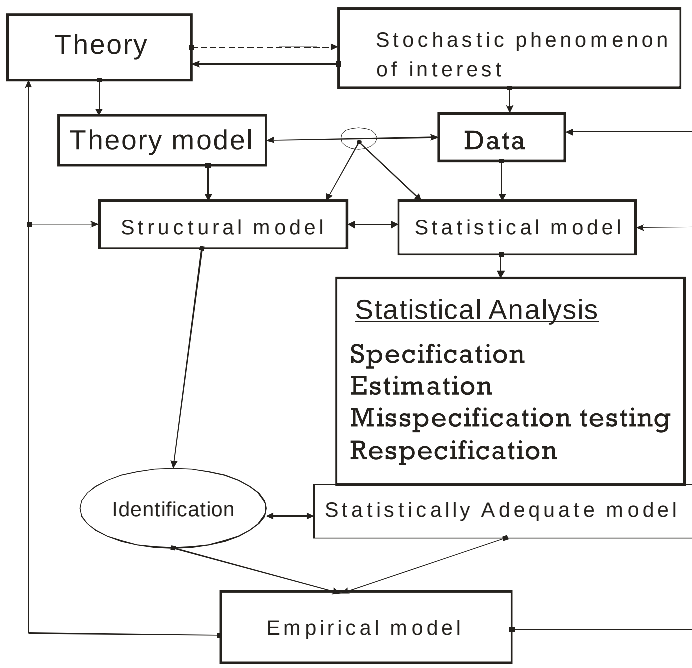

Figure 1: An empirical modeling framework

The key to unraveling the testing of assumptions argument is provided by drawing a clear distinction between *substantive* and *statistical assumptions*. The difficulty can be seen in Colander where the list of assumptions (1)-(8) invoked by Ireland (2004) is a mixture of substantive and statistical assumptions that need to be separated because their respective validity has very different implications for the reliability of inference. The substantive assumptions pertain to the realisticness issue, but the statistical assumptions pertain to the legitimate use of statistical inference as articulated by the Fisher-Neyman-Pearson approach if reliable inferences are sought. This is because when any of the statistical assumptions are invalid vis-a-vis the data in question, inferences based on the estimated model are dubious, irrespective of the substantive merits/demerits of the structural model; see Spanos (2006a-b).

Distinguishing between substantive and statistical assumptions is not as straight forward as it might seem at first sight, because it depends on having a clear-cut distinction between a structural and a statistical model where the two are inextricably bound up via the available data $\mathbf{Z}_0{:=}(\mathbf{z}_t,\ t{=}1,2,...,n)$. This is accomplished by viewing the statistical model as a purely probabilistic construal whose structure depends solely on the statistical information contained in the data. In particular, a statistical model is viewed as a parameterization of the underlying 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 in order to pose the substantive questions of interest. In this sense, a structural model is viewed as an *estimable* form of a theory model (hence, built on substantive information) in view of the available data $\mathbf{Z}_0$; see Spanos (1986, 1995).

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 (2001). 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 is:

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

and its parameters $\boldsymbol{\varphi} := (\boldsymbol{\beta}_0, \mathbf{B}_0, \mathbf{B}_1, \mathbf{B}_2, \mathbf{V})$ constitute *re-parameterization* of those of the VAR(1) model in the sense that $\boldsymbol{\varphi} = \mathbf{H}(\boldsymbol{\theta})$; see Spanos (1986).

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:	$\{\mathbf{Z}_t,\ t{\in}\mathbb{N}\}$, is a Markov process
[5]	t-invariance:	$\boldsymbol{\theta} {:=}(\mathbf{a}_0, \mathbf{A}_1, \Omega)$ are t-invariant,
	$\mathbf{a}_0 = \boldsymbol{\mu} - \mathbf{A}_1^\top \boldsymbol{\mu},\quad \mathbf{A}_1 = \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1,\quad \Omega := \boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1^\top \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1$	

It turns out that the sequence of models, theory, structural (estimable) and statistical, provides a way to foreground as well bridge the gap between theory and data; see Mayo (1996) for a similar proposal in philosophy of science. The separation is particularly crucial because statistical adequacy [the validity of the statistical assumptions vis-a-vis the data in question] is a pre-condition for any reliable statistical 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 legitimizes the use of statistical inference by securing the statistical reliability of inference:

the *actual error probabilities* are very close to the *nominal* ones. Inferences based on a statistically misspecified model are likely to lead one astray! The surest way to be led astray in drawing inferences is to apply a .05 significance test when the actual type I error is closer to .95; see Spanos and McGuirk (2001).

The notion of statistical adequacy replaces goodness-of-fit as *the sole* criterion for assessing when a fitted model 'accounts for the regularities in the data', addressing problem (ii), and at the same time bringing out the fallacious nature of (I)-(III) misleadingly invoked by the pre-eminence of theory advocates. 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. Statistical adequacy is achieved by applying thorough *Mis-Specification (M-S) testing* to probe effectively the different ways the model assumptions (e.g. [1]-[5] in table 1) might be misspecified in an attempt to confirm that, indeed, the data constitute a truly typical realization of the process defined by the model assumptions. Advocates of the pre-eminence of theory are likely to argue that M-S testing and respecification of the statistical model are just additional forms of data-mining which are also vulnerable to charges of pre-test bias. Such arguments, however, are clearly misplaced and reveal ignorance, on behalf of the critic, conerning statistical foundational issues; see Spanos (2000). Although the effectiveness of M-S testing requires judicious use of graphical techniques, there is nothing subjective about the judgment needed to validate a statistical model, and no relevant error probability depends on the intentions of the modeler; see Mayo and Spanos (2004).

It is important to emphasize that statistical misspecification and its implications cannot be fended off using locutions of the form:

> "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 for two reasons. *First*, they reveal a deep-rooted confusion between a statistical and a theory model, and *second*, it can be shown that even (seemingly) minor departures from certain probabilistic assumptions can give rise to major discrepancies between actual and nominal error probabilities, seriously undermining the associated inference; see Spanos and McGuirk (2001). Hence, in order to justify an invoked robustness one needs to verify such locutions by evaluating the discrepancy between nominal and actual error probabilities; see Spanos (2005).

One can go further and argue that by embedding a structural model into a statistically adequate model enables us to make some progress on the centuries old realisticness issue. This issue could not even be posed when a structural model cannot account for the statistical regularities in the data; this should be a minimum requirement. Beyond that one could then assess the substantive adequacy of such a structural model vis-a-vis the phenomenon of interest by probing the different ways an inference based on it could be in error; see Spanos (2006a).

To what extent can one use the methodological framework in figure 1, to shed light on the strengths and weaknesses of both perspectives, and provide the basis for a constructive dialogue between them?

The pre-eminence of theory perspective essentially ignores the right hand side, including the various statistical analysis steps that will yield a statistically adequate model which can then embed the structural model in its context in order to assess its data-acceptability. Estimating the structural model directly using the available data, by sidestepping the statistical analysis, would almost certainly give rise to completely unreliable inferences whose interpretation will be marred by numerous Duhemian ambiguities; one would not be able to apportion blame for unwarranted inferences to statistical or substantive inadequacies. No reliable inferences can be drawn concerning the adequacy of the structural model in question without first securing the statistical adequacy of the underlying statistical model.

At this stage an advocate of the pre-eminence of theory perspective is likely to ask 'what is the underlying statistical model? After all, the prespective would only specify a structural model, as the only model that matters'. It turns out that every estimable structural model has an *implicit* statistical model (often called the *reduced form*) whose statistical adequacy is a pre-condition for the statistical reliability (the actual error probabilities are close to the nominal ones) of any inference based on the estimated structural model. The relationship between the structural and (implicit) statistical models enters the discussion in relation to the problem of *identifying* the structural parameters of interest, but it's often insufficiently appreciated that such an identification is empirically vacuous [vis-a-vis the data], when the statistical model is misspecified. Moreover, the sampling theory invoked (including asymptotic Normality) for any inference based on the estimated structural model is often dubious; see Spanos (1990). Indeed, the only coherent way to assess the appropriateness of a structural model vis-a-vis data $\mathbf{Z}_0$ is to embed [parametrically nest] it into a statistically adequate statistical model and test its over-identifying restrictions. The latter restrictions carry the substantive information over and above the statistical information in data $\mathbf{Z}_0$ summarized by the statistical model. Rejection of the over-identifying restrictions provides empirical evidence against the substantive adequacy of the structural model vis-a-vis the phenomenon of interest. This operationalizes the centuries old realisticness issue by narrowing it down and rendering the structural model testable within the context of a statistically adequate model. If a structural model cannot even account for the statistical regularities in the data, it does not meet the minimum requirement of empirical adequacy, let alone substantive adequacy. Moreover, in practice one needs to go beyond the statistical significance/insignificance to substantive significance/insignificance in order to adequately address problem (iii) above, which can be achieved using severity assessments based on post-data evaluations of statistical inferences; see Mayo and Spanos (2006), Spanos (2006b).

The general-to-specific cointegrated VAR 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 (often low-level), one can proceed to specify a statistical model, like a VAR (table 1), in terms of the probabilistic structure of the underlying (vector) 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-based restrictions*, like the ones implied by *cointegration* when it exists, and attempt to relate the restricted model to certain low-level theories using notions of long-run steady-state and/or equilibrium-correction states; see Hendry (1995), Johansen (1996), Juselius (2006).

This leaves the European perspective vulnerable to the charge that their use of theoretical information is ostensible because the restricted statistical model, where the restrictions are data-based, is only tangentially connected to economic theory. In their defense, advocates of the European perspective are likely to offer a plethora empirical evidence that when the pre-eminence of theory strategy is adopted the structural models it gives rise to, like Ireland's (2004) DSGE model, are clearly rejected as empirically incongruous; see Juselius and Franchi (2007). This returns the ball back to the court of the advocates of the pre-eminence of theory perspective, with the onus on them to construct structural models which can potentially bridge the gap between theory and data. Hence, real progress in learning from data can be expected only when macro-economists are willing to face, head on, the formidable difficulties in addressing problem (i); there is no way one can circumvent that problem and expect *empirical models* that harmoniously blend both the statistical and substantive information.

4 Conclusion: reconciling the two perspectives?

The main thesis of this paper is that the methodological framework in figure 1 can be used to make suggestions for research strategies which could potentially reconcile the two perspectives.

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 purely probabilistic construal. Thorough M-S testing will determine if the latter is statistically adequate or not. Based on past experience of this author in assessing the statistical adequacy of such models, it is highly unlikely that such a model will turn out to be statistically adequate; see Juselius and Franchi (2007). This, by itself, provides empirical evidence against the structural model in question, and revisions are called for; revisions which take into consideration the statistical regularities reflected in the adequacy of the estimated statistical model.

A way to proceed might be to respecify the original statistical model with a view to reach a statistically adequate parameterization of the underlying process $\{\mathbf{Z}_t,\ t\in\mathbb{N}\}$. Respecification in this context does not refer to 'error-fixing' widely recommeded in traditional textbook econometrics, but it should take the form of assuming more appropriate probabilistic structure on $\{\mathbf{Z}_t,\ t\in\mathbb{N}\}$which is likely to give rise to more appropriate parameterizations. 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 ameliorating the original structural model by modifying it to account for 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 can be considered empirically adequate. This process might require several iterations before such a model is reached. When no statistically adequate model which would nest the structural model can be found, that in itself provides *prima facie* evidence that the latter is inappropriate or at least incomplete, and a revision is called for.

How does this strategy relate to the general-to-specific cointegrated VAR perspective? Relating the statistically adequate models reached by the latter approach to potential structural models requires one to face up to the formidable difficulties in addressing problem (i), i.e. bridging the huge gap between economic theories and the available observational data. A statistically adequate model can be used to give data a voice of its own, to reliably constrain economic theorizing, and hopefully help direct the search toward more appropriate structural models. This will provide the basis for a constructive dialogue between the two perspectives to the benefit of both. Indeed, this will hopefully provide the basis for a broader dialogue between economic theorists and econometricians that could give rise to 'learning from data'. Mayo (1996) argues that the presence of such a constructive dialogue is what distinguishes scientific from non-scientific disciplines.

What are the prospects that such a constructive dialogue will begin any time soon? I remain optimistic that the new generation of empirical economists and econometricians will eventually grow out of esteeming technical dexterity and begin to question the sanctity of the two centuries old *status quo* by seeking sounder empirical foundations for economic theories. The stranglehold that the pre-eminence of theory perspective currently has on publishing in US dominated journals can only delay the inevitable change for the better. The prospect of empirical modeling in economics losing its credibility as a serious scientific field vis-a-vis other scientists as well as policy makers, and the danger of having the same fate as Political Arithmetic toward the end of the 18th century, are very real and looming large; see Spanos (2006a, 2008b).

References

[1] Cairnes, J. E. (1875/1888), *The Character and Logical Method of Political Economy*, 2nd ed., Reprints of Economic Classics, 1965, Augustus, M. Kelley, NY.

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

[3] Colander, D. (2008), "Economists, Incentives, Judgment, and Empirical Work," Economics Discussion Papers, No 2008-12. http://www.economics-ejournal.org/economics/discussionpapers/2008-12.

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

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

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

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

[8] Hendry, D. F. (2003), "Denis Sargan and the origins of the LSE econometric methodology," *Econometric Theory*, 19: 457-480.

[9] Ireland, P. (2004), "A Method for taking Models to the Data" *Journal of Economic Dynamics and Control,* 28: 1205-1226.

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

[11] Johansen, S. (2006), "Confronting the Economic Model with the Data," pp. 287-300 in D. Colander, *Post Walrasian Macroeconomics: Beyond the DSGE Model*, Cambridge University Press, Cambridge.

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

[13] 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.

[14] Kennedy, P. (2008), *A Guide to Econometrics*, 6th ed., Blackwell Publisher, Oxford.

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

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

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

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

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

[20] 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.

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

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

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

[24] 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.

[25] Sargent, T. (1976), "A Classical Macroeconomic Model for the United States," *Journal of Political Economy*, 84: 207-238.

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

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

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

[29] Spanos, A. (1999), *Probability Theory and Statistical Inference: Econometric Modeling with Observational Data*, Cambridge University Press, Cambridge.

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

[31] Spanos, A. (2001), "Time Series and Dynamic Models," pp. 585-609, *A Companion to Theoretical Econometrics*, ed. by B. Baltagi, Blackwell, Oxford.

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

[33] 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.

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

[35] Spanos, A. (2008a), "Theory Testing in Economics and the Error Statistical Perspective," forthcoming in *Error and Inference*, edited by D. G. Mayo and A. Spanos, Cambridge University Press, Cambridge.

[36] Spanos, A. (2008b), "Statistics and Economics," in the *New Palgrave Dictionary of Economics*, 2nd edition, Eds. S. N. Durlauf and L. E. Blume. Palgrave Macmillan, London.

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