17 On the role of theory and evidence in macroeconomics

Katarina Juselius[1]

17.1 INTRODUCTION

Economists frequently formulate an *economically* well-specified model as the empirical model and apply statistical methods to estimate its parameters. In contrast, statisticians might formulate a statistically well-specified model for the data and analyze the statistical model to answer the economic questions of interest. In the first case, statistics are used passively as a tool to get some desired estimates, and in the second case, the statistical model is taken seriously and used actively as a means of analyzing the underlying generating process of the phenomenon in question.

The general principle of analyzing statistical models instead of applying methods can be traced back to R.A. Fisher. It was introduced into econometrics by Trygve Haavelmo (1944) and operationalized and further developed by Hendry and Mizon (1993), Hendry (1987), Johansen (1995), Juselius (2006), and followers. Haavelmo's influence on modern econometrics has been discussed, for example, in Hendry et al. (1989) and Andersen (1991).

Because few observed macroeconomic variables can be assumed to be fixed or predetermined a priori, Haavelmo's approach to econometrics requires a probability formulation of the full process that generated the data. Thus, the statistical model has to be based on a full system of equations. The computational complexities involved in the solution of such a system were clearly prohibitive at the time of Haavelmo's monograph when even the estimation of a multiple regression was a non-trivial task. In today's computerized world, it is certainly technically feasible to adopt Haavelmo's guidelines to empirical econometrics. Although the technical difficulties have been solved long ago, most articles in empirical macroeconomics do not seem to follow Haavelmo's general principles, despite being stated very clearly in his monograph.

I shall argue here that the cointegrated vector autoregressive (hereafter CVAR) approach offers a number of advantages as a general framework for addressing empirical questions in (macro)economics without violating Haavelmo's general probability principle. I shall also argue that the

CVAR approach is likely to meet Summers's 1991 critique (Summers, 1991) on many points. He claimed in his article titled 'The Scientific Illusion in Empirical Macroeconomics' that empirical economics has exerted little influence on the development of economic theory and provided little new insight on economic mechanisms. As an illustration, he mentioned empirical analysis of representative agent models, which involves estimating a few deep parameters characterizing preferences and technology. He contrasted this analysis with sophisticated statistical techniques that impose minimal theoretical constraints on the data, as exemplified by a vector autoregressive (VAR) model à la Sims. Summers argued that both approaches have been unable to explain the inherently richer and more complicated macroeconomic reality. He concluded that less formal examination of so-called stylized facts (correlations, mean growth rates, graphs) has resulted in more fruitful economic research. Almost 20 years have passed since Summer's critique and this seems a good opportunity to revisit his main arguments. Has the marriage between theory and evidence improved? Has empirical economics become more enlightening?

One of the key developments after Summers's critique was the so called dynamic stochastic general equilibrium (DSGE) modeling of the macroeconomy. These models combine the assumptions of a representative agent and the 'Rational expectations hypothesis' (REH) together with a dynamic stochastic structure à la Sims's VAR. The added dynamics and stochasticity render the models more flexible than earlier representative agent models. Nonetheless, in the way that they are practiced, one may say that the DSGE approach gives the primary role to the theory model and a subordinate role to the VAR. In this sense the DSGE model embodies the principle of pre-eminence of theory over empirics or what I shall call the 'theory-first' approach. Hardly surprisingly, not everyone in the profession was prepared to adopt the theory-first approach exemplified by the DSGE models as the best way to describe and understand our complex empirical reality. Many methodologically oriented scholars were vigorously debating the role of theory and empirical evidence in macroeconomics as illustrated by the discussions in Backhouse and Salanti (2000a, 2000b) and Colander (1996, 2000, 2001, 2006). These debates however, seemed, only to have a marginal (if any) effect on the empirical practise of the economics profession. At the beginning of the new century the popularity of the DSGE models in graduate programs and among editors of top economics journals and researchers at central banks suggested that economics as a science had finally converged to a state of unanimity regarding both theory and how to apply it to data.

This view was challenged when the crisis struck with a suddenness that took most economists, central bankers and politicians by complete surprise. Obviously, standard models must have lacked important features as they failed to warn their users about the growing vulnerability of the macroeconomy to the crisis. As the crisis accelerated, the chorus of critical voices became a tsunami that washed over newspapers, blogs, and so on. The debate on how to do economics was again fresh and alive. See for example discussions in Colander et al. (2008, 2009), Katlesky (2009), Lejonhufvud (2009), Krugman (2009), Cochrane (2009), Juselius (2009a, 2011a) and Lawson (2009). Summers's critique of the scientific value of empirical models in economics seems as relevant as it was 20 years ago.

I shall here question the pre-eminence of theory over empirics in economics and argue that empirical econometrics needs to be given a more important and independent role in economic analysis, not only in order to have some confidence in the soundness of our empirical inferences, but also to uncover empirical regularities that can serve as a basis for new economic thinking.

In the first part of this chapter I contrast the theory-first approach with what I call the data-first approach as a way of bridging economic theory and empirical evidence. To fix ideas, I make use of the article 'Taking a DSGE Model to the Data Meaningfully' (Juselius and Franchi, 2007), which provides a detailed discussion of 'A Method for Taking Models to the Data' (Ireland, 2004a). Since the discussion can be seen as a critical attack on Ireland's article, I would like to stress that it was chosen because it was well documented, all results are reproducible, and it uses sophisticated econometrics.

I then argue that one important reason why the DSGE theory-first approach does not seem to have resolved Summers's critique is that it does not properly account for the pronounced non-stationarity, such as unit root persistence and breaks, typical of economic data, neither from a theoretical nor an econometric point of view. To have some confidence in the soundness of the inference being made, the number of stochastic trends has to be tested, not just assumed; parameter constancy has to be checked, not just assumed; dynamics have to properly fit the data, and so on. I demonstrate that these features can be modelled and tested with the (cointegrated) VAR model and argue that the non-stationarity of the data has strong implications for how to model expectations and for the *ceteris paribus* clause in economic models. This data-first approach shows how economic theory can be confronted with the basic regularities in the data, thereby exposing lack of empirical relevance and sometimes even inconsistencies.

In the second part of the chapter, I discuss how a statistically adequate VAR analysis often delivers additional (often theoretically puzzling) results. If these are taken seriously, they can generate new economic hypotheses that subsequently can be tested on new data. Thus, a VAR analysis based on full information maximum likelihood analysis of the data can be used to test prior hypotheses and to generate new ones, thereby replacing 'simple stylized facts' such as correlations, graphs, and so on with more sophisticated facts that better describe the non-stationary world in which individual decisions and market outcomes unfold. By doing so, it has the potential of providing a basis for new economic thinking.

17.2 THEORY-FIRST VERSUS DATA-FIRST

The basic dilemma of empirical macro modeling is that the reality behind the available macroeconomic data is so much more rich and complex than the often narrowly analyzed problem being modeled by the theory. How to treat these 'additional' features of the data, which often go against the *ceteris paribus* assumptions of the economic model, has divided the profession into the proponents of the so called specific-to-general versus general-to-specific approach to empirical economics. To emphasize my subsequent arguments I shall here call them theory-first versus data-first. For a detailed methodological discussion of the two approaches, see for example Gilbert (1986), Hendry (1995, 2009), Juselius (2006), Pagan (1987) and Spanos (1995, 2006, 2009).

The former, more conventional, approach starts from a mathematical formulation of a theoretical model and then expands the model by adding stochastic components. The aim is to estimate the parameters of a 'stylized' economic model, while ignoring the wider circumstances under which the data were generated. These factors are then dumped into the residual term, causing its variance to be large. This practice has important implications for the power of empirical testing, often leading to a low ability to reject a theory model even when it provides a poor description of the basic regularities in the data. As a result, competing theory models are often not rejected despite being tested against the same data. See Cooley and Leroy (1985) for an early discussion. Furthermore, the statistical inference in such models is usually based on a number of untested and empirically questionable assumptions so the 'significance' of estimated parameters may lack scientific meaning.

This approach is strongly influenced by the legacy of the pre-eminence of theory, which presumes that the basic economic mechanisms can be pre-specified, that is, we know which variables are exogenous, which are adjusting when the system has been pushed out of steady state by exogenous shocks, how various interventions have affected the system, ans so on. Econometrics in this case play the subordinate role of 'quantifying' theoretically meaningful parameters assumed to be empirically relevant on a priori grounds. This approach can only be defended if the probabilistic assumptions comprising the underlying statistical model are satisfied vis-à-vis the data in question (Spanos, 2009). If not, it produces empirically irrelevant and possibly misleading results.

The data-first approach starts from an explicit stochastic formulation of all data (selected on the basis of the theory model in question) and then reduces the general statistical model by simplification testing. It answers the economic questions of interest by imbedding the economic model within the statistical model and uses strict statistical principles as a criterion for a good empirical model. In this case, the statistical model ties economic theory to the data when it nests both the data-generating process and the theoretical model, so that the parameters of the theoretical model can be derived from the parameters of the statistical model (Hoover et al., 2008).

It recognizes from the outset the weak link between the theory model and the observed reality. For instance few, if any, theory models fully allow for basic characteristics of macroeconomic data such as path-dependence, unit-root non-stationarity, structural breaks, shifts in equilibrium means, and location shifts in general growth rates. Such empirical features of economic data are generally at odds with the prevailing theory-first paradigm, which assumes a few constant structural parameters describing technology and preferences, REH and some ad hoc dynamics. In contrast, the data-first methodology works by allowing the empirical regularities in data to speak as freely as possible about underlying theoretical relationships, thereby allowing us to discriminate between empirically relevant and irrelevant theories, but also to discover new evidence for which prior hypotheses have not yet been formulated. All this will be discussed in the next sections using Ireland (2004a) as an illustrative case study.

17.3 IRELAND'S DSGE MODEL: A FIRST CHECK

The hypothesis that aggregate technology shocks alone drive business cycle fluctuations is a key feature of the real business cycle (RBC) model discussed in Ireland (2004a). More specifically, the model assumes a utility-maximizing representative agent, who chooses between consumption and total hours worked in an economy subject to a constant returns to scale technology described by a Cobb–Douglas production function. The latter is defined for capital and labor subject to labor-augmented technological progress, which is assumed to have grown with a constant rate over the examined period 1948–2002. Capital and total factor productivity are assumed to be unobserved and generated from cumulated identical shocks. The model describes a closed economy without a government sector.

To take the highly non-linear RBC model to the data, Ireland applies log-linearization around theoretical steady-state values making use of the following basic assumptions: (1) total factor productivity follows a first-order stationary autoregressive model; and (2) the growth rate of labor-augmenting technological progress is constant over time. The implications of these assumptions are that: (1) trend-adjusted output, consumption, investment and capital are stationary around their steady-state values; (2) the rate of labor-augmenting technological progress is constant and results in identical linear trends in output, consumption, investment and capital; and (3) technology shocks are pushing both total factor productivity (TFP) and capital. The model's three equations (output, consumption and labor) are made 'more flexible' by adding a VAR(1) process. Most model parameters are estimated using a Kalman filter non-linear optimization routine that can handle the singularity of the model.[2] The exceptions are the parameters for depreciation rate and the discount rate, which are calibrated. The estimated structural parameters are constrained to satisfy the restrictions implied by theory.

The reported estimates are claimed to be maximum likelihood (ML) estimates. But ML requires that the assumptions underlying the model are correct. Thus, they have to be checked, not just assumed. In this context it is useful to draw a distinction between substantive and statistical assumptions (Spanos, 2009) as their respective validity has different implications for inference. The substantive assumptions pertain to the centuries-old issue of the realism of economic theories, whereas the statistical assumptions pertain to the reliability of the statistical inference. Because the Ireland article did not report tests of these assumptions, Juselius and Franchi (2007) first replicated all results reported in Ireland's article and then performed a detailed check of the statistical and some of the substantive assumptions.

Parameter constancy is one of the substantive assumptions: if parameters are structural they ought to remain constant despite changes in policy regimes. Over the sample period, the US economy experienced a major policy regime shift around 1979. In a footnote, Ireland mentions that parameter constancy before and after this date has been tested and rejected. Nonetheless he disregarded this evidence, because parameter estimates were 'quite similar' before and after that date.

The stationarity assumption needed for the log-linearization around the constant steady states can be assessed based on estimates of the characteristic roots of the model, of which the largest were 0.9987, 0.94 and 0.88[3]. A root of 0.9987 is admittedly less than one, but in practice indistinguishable from a unit root. Even a root of 0.94 is pretty close to unity, suggesting an additional source of pronounced persistence in the data.

One consequence of the assumptions of identical deterministic growth rates for output, consumption and capital, when the actual rates differed markedly, and stationarity, when there are near unit roots in the data, can be inferred from Figure 17.1.[4] The deviations from the assumed steady-state values are very persistent and systematically either positive or negative. Except for hours worked, they never cross the zero line over a period of 50 years.

The statistical misspecification tests of Ireland's model showed that the null hypotheses of residual normality, no autocorrelation and no autoregressive conditional heteroskedasticity (ARCH) were rejected for essentially all variables. Furthermore, the cross correlogram showed significant correlations between the errors which are assumed independent. Thus, the first set of diagnostic tests revealed clear violations of the distributional assumptions. Hence, the probability model is not correctly specified; the reported statistical inference is not maximum likelihood and the p-values calculated from standard normal distributions (t, F, and χ^2) may be completely unreliable. In particular, the assumption that data are stationary, when in fact they are very close to being non-stationary, is likely to make all inference on steady-state values meaningless (see Johansen, 2006).

There is, however, an easy solution to the problem of unit roots in the data and its effect on inference. By transforming the data into stationary components using differencing and cointegration (see, for example, Engle and Granger, 1987; Hendry, 1987, 2009; Johansen, 1988, 1995; Juselius, 2006) standard inference would apply again. This is also the preferred solution in Ireland (2004b). That the assumed probability model was not correctly specified does not as such imply that the assumed RBC model is incorrect. Therefore, Juselius and Franchi (2007) examine whether the RBC hypothesis is a good description of US business cycles by recasting the basic hypotheses in the Ireland model as statistical tests in a well-specified cointegrated VAR (CVAR) model.[5]

17.4 THE RBC HYPOTHESIS: A CVAR CHECK

17.4.1 A Theory-Consistent CVAR Scenario

To test the basic RBC assumptions underlying Ireland's DSGE model formulation, we need to formulate all basic (substantive and statistical) assumptions as testable hypotheses on the VAR model. Such a theory-consistent VAR scenario can be seen as a bridge between the theory model

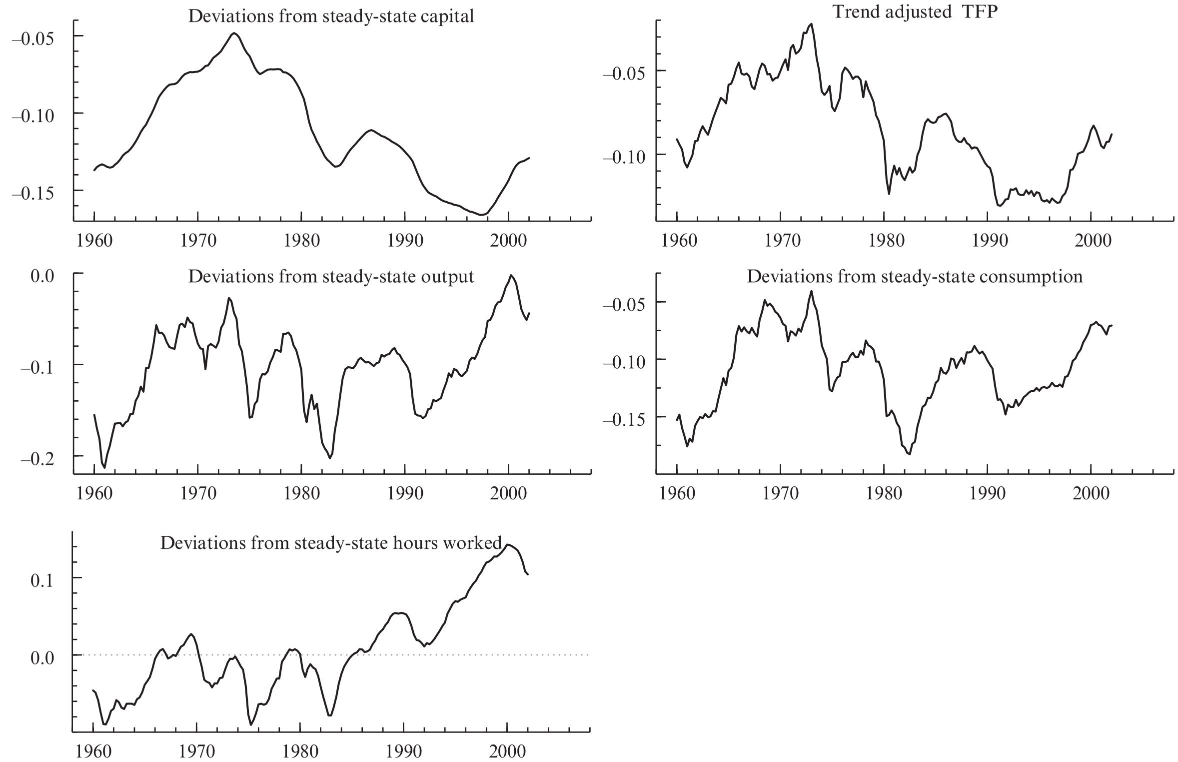

Figure 17.1 Deviations from trend-adjusted steady-state values for capital, income, consumption, labor and TFP

and the VAR and is a summary of necessary conditions on the data for the model to be empirically relevant.

One of the roots (associated with total factor productivity) was very close to unity and was approximated with a unit root, that is, we assume that total factor productivity, a_t, is a unit root stochastic process instead of an AR(1) with a near unit root:

$$a_t = a_{t-1} + \varepsilon_t = \sum_{i=1}^{t} \varepsilon_i + a_0, \quad (17.1)$$

where ε_t is a permanent TFP shock. The cumulation of these shocks, $\sum_{i=1}^{t}\varepsilon_i$, is called a (first-order) stochastic trend and describes the development of total factor productivity over time. A variable containing a first-order stochastic trend is called integrated of order one, in short I(1). The difference between a first-order stochastic and a deterministic linear trend is that the permanent increments of a stochastic trend change randomly, whereas those of a deterministic trend are constant over time.

When two variables share the same cumulated permanent shocks we say that the two variables are cointegrated. For example, if capital and income share the same TFP stochastic trend (17.1) it implies that they have followed the same long-run stochastic growth path and are therefore likely to be causally associated. In a growing economy, many variables would contain a linear deterministic trend. But as the source of the linear increment is generally not identified, the fact that two variables share a linear deterministic trend does not imply that they are causally related (Yule, 1926). Therefore, finding significant cointegration is a much more demanding and relevant measure of long-run association than a significant regression relationship. For an expository introduction to integration and cointegration see Hendry and Juselius (1999, 2000).

Furthermore, an important advantage of classifying data according to their order of integration, I(0), I(1) and I(2)[6] is that it allows us to classify variables and relations according to their 'persistency profiles', that is, we can discriminate between relations which exhibit pronounced persistence and relations which do not (Juselius, 2010, 2011b). For example, the strong persistence visible in the graphs of the steady-state errors in Figure 17.1 suggests that: (1) the long-run path of Ireland's data is not adequately described by a common deterministic trend; and (2) the steady-state errors contain unexplained persistence that suggests serious misspecification.

Assuming one stochastic trend and three cointegration relations, a theory-consistent CVAR scenario for Ireland's model is given by (17.2) and (17.3). It is useful to start with the formulation of the pushing forces:

$$\begin{bmatrix} y_t \\ c_t \\ h_t \\ k_t \end{bmatrix} = \begin{bmatrix} d_1 \\ d_1 \\ 0 \\ d_1 \end{bmatrix} [a_t] + \begin{bmatrix} b_1 \\ b_1 \\ 0 \\ b_1 \end{bmatrix} [t] + \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \\ v_{4,t} \end{bmatrix} \tag{17.2}$$

where y_t is the log of output, c_t is the log of consumption, h_t is the log of hours worked, k_t is the log of capital, $a_t = \sum_{i=1}^{t} \varepsilon_i$ is the sum of total factor productivity shocks and v_t is a stationary moving average (MA) process. It implies: (1) a non-stationary Cobb–Douglas production function, $y_t - \theta k_t - (1-\theta)h_t$ is $I(1)$; (2) a stationary consumption–income ratio, $c_t - y_t$ is $I(0)$, (3) a stationary capital–income ratio, $k_t - y_t$, is $I(0)$; and (4) that hours worked is stationary, h_t is $I(0)$. Thus, under scenario (17.2) the three cointegration relations should correspond to the relations in (2)–(4).[7] According to Ireland's model, shocks to technical progress should explain the development over time in TFP and capital, whereas output, consumption and labor would be adjusting (that is, endogenous). The latter can be formulated in the following equilibrium error correcting model:

$$\begin{bmatrix} \Delta y_t \\ \Delta c_t \\ \Delta h_t \\ \Delta k_t \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \alpha_{31} \\ \alpha_{12} & \alpha_{22} & \alpha_{32} \\ \alpha_{13} & \alpha_{23} & \alpha_{33} \\ \alpha_{14} & \alpha_{24} & \alpha_{34} \end{bmatrix} \begin{bmatrix} (c - y - \mu_{01})_{t-1} \\ (k - y - \mu_{02})_{t-1} \\ (h - \mu_{03})_{t-1} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_1 \\ 0 \\ b_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \end{bmatrix} \tag{17.3}$$

The scenarios (17.2) and (17.3) comprise a set of necessary conditions that need to be supported by the data for the Ireland RBC model to pass a first test of empirical relevance.

17.4.2 A CVAR Reality Check of Ireland's Model

An unrestricted VAR(2) model in levels was first estimated and tested for misspecification, and then revised accordingly. The null of constant parameters for the periods 1960:1–1979:4 and 1981:2–2002:1 was, however, strongly rejected: the assumption that the structural parameters are constant over time did not seem tenable with the information in the data. The VAR analysis was, therefore, done separately for the two periods and produced the following general results: (1) The first period was fairly well described by the estimated model, but the second period less so;[8] (2) in both periods there seemed to be at least two stochastic trends (rather than one), one of which was associated with permanent shocks to consumption and the other to labor; (3) a trend-stationary Cobb–Douglas production function with plausible coefficients seemed to work well in the first, but not in the second period; (4) hours worked was found to be non-stationary against the stationarity assumption in Ireland's model.

Altogether, the results suggested that total factor productivity and technological progress might have been well approximated by a linear trend in the first period but not in the second, and that it was demand shocks rather than shocks to TFP that have triggered off US business cycles. Thus, the information in the data seemed more consistent with a Keynesian than an RBC explanation of US business cycles. When this is said, the CVAR results should only be considered tentative, as a better model specification is clearly needed. This is particularly so for the second period where the results were based on an econometrically unsatisfactory model.

Thus, the conclusion reached in Ireland's article (that the real business cycle theory model is able to explain the long business cycles in the US post-war period) was not based on correct statistical inference and the conclusion was reversed when based on a well-specified CVAR model. In spite of the sophisticated dynamic and stochastic specification, Ireland's DSGE model seemed to lack important features needed to account properly for the complexity of the economic reality: The basic predictions from Ireland's DSGE model fell outside the confidence bands of the empirical reality.

17.4.3 Pulling Sophisticated Empirical Facts out of Data

Summers (1991) concluded his critique by claiming that a less formal examination of empirical facts has generally resulted in more fruitful economic research. While possibly useful, such stylized facts are nevertheless too coarse and can even be misleading when data are non-stationary, because simple correlation coefficients and mean growth rates are only well defined for stationary, but not for non-stationary data (Yule, 1926).

The unrestricted VAR model, properly specified and tested, can be seen as a convenient summary of the covariances of the data (Hendry and Mizon, 1993; Juselius, 2006, Ch. 3) and, therefore, represents the reality we would like our theoretical model to describe. But, because the VAR is heavily overparametrized, it is not very informative as such. Its usefulness derives from its ability to reduce the number of parameters until further restrictions significantly change the value of the likelihood function. If correctly done, the final parsimonious VAR model would describe regularities in the data without suppressing any relevant information and, therefore, would provide a set of 'statistical regularities' that the theoretical model should explain in order to be empirically relevant.

In particular, by combining differenced and cointegrated data, the CVAR model offers a natural way of analyzing economic data as short-run variations around moving long-run equilibria. Longer-run forces are themselves divided into the forces that move the equilibria (pushing forces, which give rise to stochastic trends) and forces that correct deviations from equilibrium (pulling forces, which give rise cointegrating relations). Interpreted in this way, the CVAR has a good chance of nesting a multivariate, path-dependent data-generating process and relevant dynamic macroeconomic theories. One could say that the CVAR model gives the data a rich context in which they are allowed to speak freely (Hoover et al., 2008). See also Møller (2008) for a detailed exposition of how to translate basic concepts of macroeconomic models into testable concepts of the CVAR model.

The CVAR analysis of Ireland's DSGE model illustrates this latter point. When data were allowed to speak freely, they spoke about a number of empirical regularities that had been silenced by prior restrictions: the number of stochastic trends were two or three in the data, but one in Ireland's model; the two driving forces originated from shocks to consumption and labor in the data, but from TFP and capital in Ireland's model; the lag order of the VAR was two in the data, but one in Ireland's model; the variables were trend non-stationary in the data, but trend-stationary in Ireland's model; there was strong evidence of structural breaks in the data, whereas a constant DSGE structure was assumed in Ireland's model.

An advantage of analyzing the Ireland data separately for the period before and after 1980 was that it gave an opportunity to compare similarities and differences between the two periods and, therefore, get a first idea of which economic mechanisms had changed and why. In particular, it was possible to demonstrate that income, consumption, labor and capital did a reasonable job in 'explaining' business cycle movements in the first period (though with a Keynesian flavor), but a much less satisfactory one in the more recent period indicating that some important information is missing in the theoretical set-up. Section 17.6 will argue that it is the world wide deregulation of financial markets that has influenced US savings and investment decisions and this information is missing in the present set-up.

17.5 ECONOMIC MODELING AND NON-STATIONARY DATA

The lack of empirical support may not come as a big surprise to economists, who would argue that their models are not meant to be close approximations to the economic reality. Many would argue that by adding new features to these models they will gradually improve.[9] I would like to challenge this view and argue that the strong evidence of non-stationarity in data points to a more fundamental problem with many economic models which are intrinsically developed for a stationary world despite containing some added-on persistence and dynamics. As will be discussed below, non-stationarity has strong implications for important aspects of economic models, such as expectations, risk versus uncertainty, and the robustness of policy conclusions to the empirical validity of the *ceteris paribus* clause.

17.5.1 The Role of Expectations in a Non-stationary World

The strong evidence of (near) unit roots and (structural) breaks in economic variables and relations suggest that economic behavior is often unpredictable. In such a non-stationary world Clements and Hendry (1999, 2008) show that forecasts from constant-parameter theory models, assumed to be correct from the outset, are likely to perform poorly. Since rational expectations models imagine economic agents who are able recursively to foresee future outcomes with known probabilities, they are inconsistent with the prevalence of structural breaks in the data (Hendry and Mizon, 2011). Indeed, these models have often often failed to describe macroeconomic data satisfactorily. The strong prevalence of non-stationarity in economic time series, in itself, is evidence that we do not know in which direction the future is moving. To act as if we do would indeed be highly irrational. In the words of Keynes (1937):

> By 'uncertain' knowledge, let me explain, I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty . . . The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence . . . About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know.

But, if rational expectations have to go, what should replace them? Frydman and Goldberg (2007, 2008, 2011) provide an alternative approach dubbed 'Imperfect knowledge economics' (IKE), which recognizes that the process driving outcomes in modern economies changes at times and in ways that no one can specify ahead of time up to a random error. Such change arises in part because individuals's forecasting strategies, which play an important role in driving market outcomes, change in ways that they themselves, let alone an economist, cannot fully pre-specify.

IKE holds out the possibility that although change in modern economies does not unfold mechanically, it may exhibit qualitative regularities that can be modeled both theoretically and empirically. Because IKE models impose qualitative restrictions on change, it generates qualitative implications which, though testable, are looser than those derived under REH. Johansen et al. (2010), Juselius (2010), and Frydman et al. (2008) find that the implications of REH-based models of nominal and real exchange rates are strongly rejected in favor of IKE-based models.

One of the key implications of Frydman and Goldberg's (2007, 2008, 2009) IKE model of asset markets is that imperfect knowledge can lead to greater persistence in asset price fluctuations than implied by REH-based models. Thus the IKE model is able to explain the long swings in asset prices away from and toward benchmark levels that characterize asset markets.

17.5.2 The *Ceteris Paribus* Clause and Non-stationarity

It is a common practise to simplify a theory model by using the *ceteris paribus* clause, 'everything else unchanged'. However, the empirical relevance of the *ceteris paribus* assumption in a theory model is likely to be strongly affected by the order of integration of the *ceteris paribus* variables. If they are stationary, the conclusions are more likely to remain robust than if they are non-stationary. This is because a non-stationary *ceteris paribus* variable if included in the empirical model is likely to influence the cointegration results and, therefore, the conclusions of the model's steady-state behavior. Since no variables can be kept artificially fixed in the real economy, the empirical problem needs to be addressed in the context of 'everything else changing' and the impact of the *ceteris paribus* variables needs to be brought into the analysis by conditioning.

Consider, for example, the empirical CVAR analysis of Ireland's RBC model discussed in section 17.4.2. In particular, it demonstrated that the first period up to 1979 was reasonably well described by two pulling and two pushing forces, though with a Keynesian rather than a RBC explanation of US business cycles. For the more recent period the chosen information set was not sufficient: there was only weak evidence of cointegration, suggesting that some important determinants are missing. This can be illustrated with the graphs of the first two cointegration relations, the first

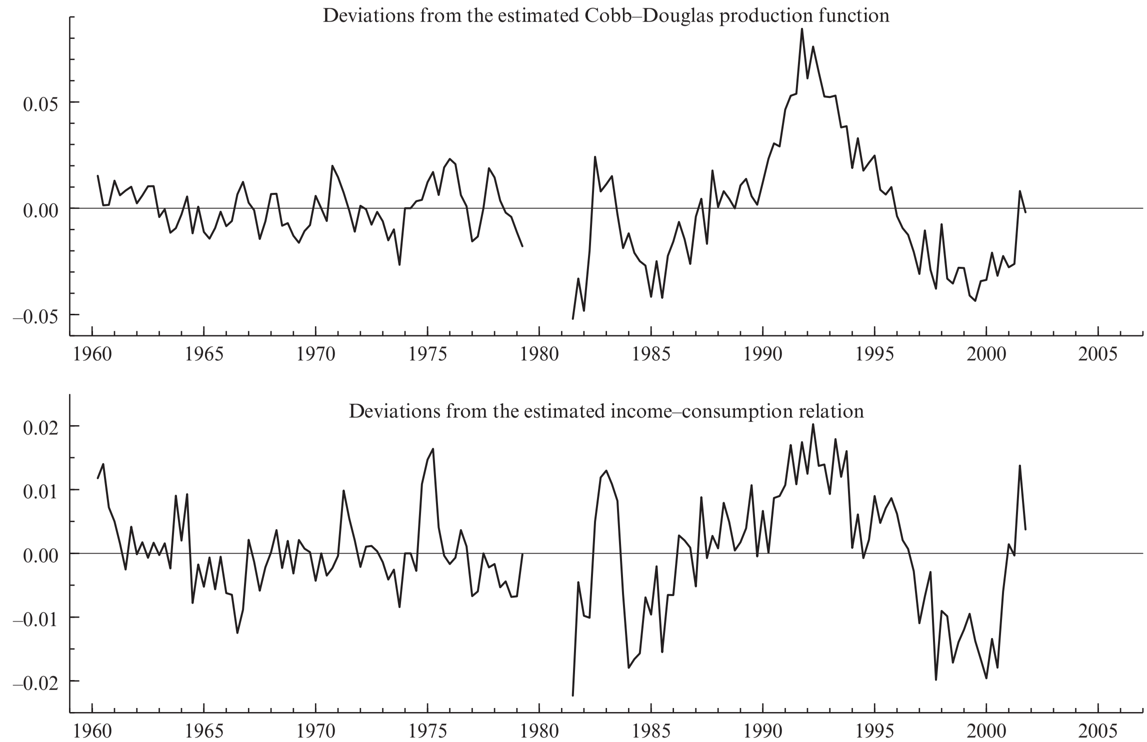

Figure 17.2 Deviations from the estimated Cobb–Douglas function (upper panel) and the estimated income–consumption relation (lower panel) over the two regimes

one describing deviations from a homogeneous Cobb–Douglas relation, the other from the consumption–income ratio. There is a striking difference between the two periods: In the first period, deviations from fundamental long-run values are reasonably modest; in the second, these are characterized by long and persistent swings.

Because a major difference between the two periods is the deregulation of financial markets in the mid-1980s, a first hypothesis that springs to mind is the importance of long swings in asset prices. This would be consistent with the IKE theory discussed above as well as with the notion of reflexivity (Soros, 1987) suggesting that such pronounced movements away from long-run benchmark values may strongly affect aggregate behavior in the real economy and vice versa.

However, if the information set is increased by new variables, for example by financial and housing wealth, reflecting booms and busts of global financial markets, previous empirical conclusions may not remain unchanged. For example, in the enlarged model it may very well be empirical shocks to financial or housing wealth that are now pushing the system, rather than shocks to consumption and hours worked. Therefore, equating a residual with an autonomous shock, as is frequently done, can be totally misleading unless the model contains all relevant variables.[10]

Such concerns are, however, easily met as it is straightforward to increase the information set, building on previously obtained results (the cointegration property is invariant to changes in the information set) thereby learning systematically how sensitive previous conclusions are to the *ceteris paribus* clause. Thus, the CVAR story is not claimed to be 'structural' in the usual sense, albeit it often has a lot of empirical content. This is contrary to Ireland's DSGE analysis which tells a 'structural' story, but with little empirical content.

17.6 'SHERLOCK HOLMES' ECONOMETRICS

More than a quarter of a century ago, when Søren Johansen and I started working on the CVAR approach, I was taken by the beauty of this model, its rich structures and its potential for addressing highly relevant questions within a stringent statistical framework. What I did not expect was that the data consistently refused to tell the stories they were supposed to. After many frustrating attempts, it seemed that I had the choice of either forcing data to tell a theoretically acceptable story or approaching the complex reality without the guide of a reliable theory. I chose the latter and started using the cointegrated VAR model in the spirit of Sherlock Holmes, an experience that best can be described as a long sequence of 'why's'.

Though economic puzzles are probably harder to solve than crimes, I believe it is in the Sherlock Holmesian spirit that a well-structured empirical CVAR analysis can inspire new economic thinking. It is based on the following important principles: (1) data are allowed to speak as freely as possible against a background of not just one but several theories; (2) falsification is considered more important than verification; and (3) results that go against conventional wisdom are considered more interesting than confirmatory results.

It has a Marshallian rather than a Walrasian flavour, expressed by Kevin Hoover as

> The Walrasian approach is totalizing. Theory comes first . . . The Marshallian approach is archaeological. We have some clues that a systematic structure lies behind the complexities of economic reality. The problem is how to lay this structure bare. To dig down to find the foundations, modifying and adapting our theoretical understanding as new facts accumulate, becoming ever more confident in our grasp of the super structure, but never quite sure that we have reached the lowest level of the structure. (Hoover, 2006)

As an illustration, I shall give a brief account of such a Sherlock Holmesian CVAR process and discuss how it has generated new hypotheses that may inspire new economic thinking.

17.6.1 Puzzling Features in the Data

Based on standard economic models it is often difficult to explain all the pronounced persistence in economic data, in particular such persistence that seems to originate from other sources than those associated with shocks to technology, preferences and money supply. Also structural breaks are common in economic data but are puzzling viewed from standard REH-based theoretical models. Such breaks can sometimes be handled econometrically by equilibrium mean shifts or changes in mean growth rates but, as Ireland's RBC model analysis illustrated, can also be of such fundamental character that a split sample analysis is required.

What do we find when analyzing data using the CVAR in the spirit of Sherlock Holmes? A robust finding is that an important regime shift seems to have occurred at the beginning of the 1980s. Few models of the macroeconomy are able to pass statistical tests of parameter constancy for the periods before and after 1980. The the main reason is that the division into exogenous and endogenous forces seems to have undergone a change: A new type of persistence influencing the number and the origin of the stochastic trends seems to influence the data after this date.

Most economists would believe that the change in US monetary policy that was the reason for the change. But data consistently refused to tell such a story. However, digging down into the data gave some other clues: the real exchange rates, the real interest rates, and the term spreads were exhibiting a pronounced persistence that seemed untenable with stationarity, as illustrated by the graphs in Figures 17.3, 17.4 and 17.5.[11] The inability to reject the unit root hypothesis suggested that many basic parity conditions, such as the purchasing power parity (PPP), the uncovered interest rate parity, the domestic and international Fisher parity and the 'expectations hypothesis' for the term spread were not working as standard REH-based theory would predict. As these variables can be associated with financial market behavior, it seemed plausible that this persistence was associated with the worldwide deregulation of financial markets. The hypothesis that the 'new' puzzling persistence in our macro models is due to speculative behavior in financial markets was suggested by Sherlock Holmesean VAR analyses some 20 years ago.

17.6.2 Empirical Regularities in the Financial Markets

As the recent and previous financial crises have demonstrated, the boom-and-bust behavior in the stock market and the market for foreign exchange is likely to have a strong impact on the real sector of the economy. Such behavior is hard to explain with REH-based models and has, therefore, often been considered puzzling. In particular, the long swings behavior of real exchange rates under currency floats has puzzled economists.12 Figure 17.3, upper panel, illustrates why: the nominal exchange rate moves in long persistent swings around relative prices. That nominal exchange rates exhibit long swings in periods of currency float, whereas relative prices do not, explains why nominal and real exchange rates tend to resemble each other.

Such persistent movements away from competitive long-run fundamental values are almost bound to have a strong impact on price competitiveness among tradables. In the absence of equilibrium-correcting nominal exchange rates, one would, therefore, expect goods prices to equilibrium-correct. But the high variability with which nominal exchange rates move prevents prices from taking the full burden of adjustment. Therefore, to restore equilibrium in the goods market, nominal interest rates would have to compensate for these persistent swings in the real exchange rates, producing a similar persistence in nominal interest rates. The latter is illustrated in the upper panel of Figure 17.4, which exhibits long and persistent swings in the US–German long-term interest rate differential. Similar persistent swings can be shown for short-term interest rate differentials, albeit with greater variability. Figure 17.5 demonstrates that also the short-long

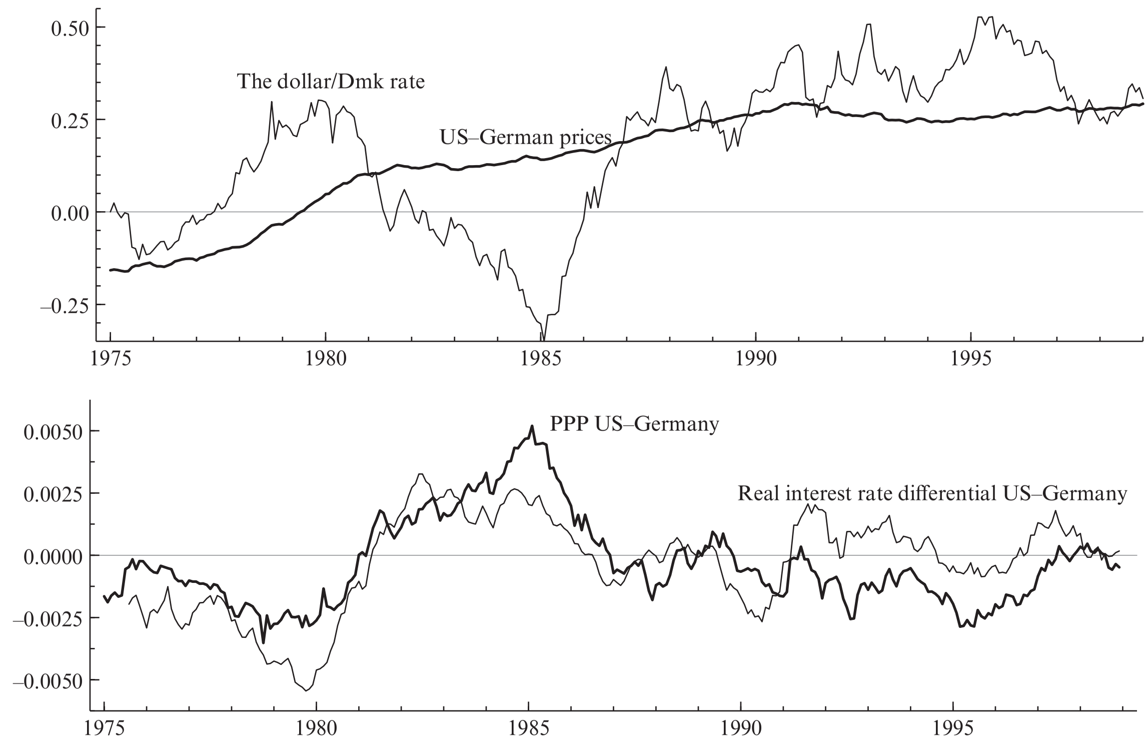

Figure 17.3 *Mean adjusted relative prices and the nominal exchange rate (upper panel) and the real exchange rate and the real interest rate differential (lower panel) between US and Germany*

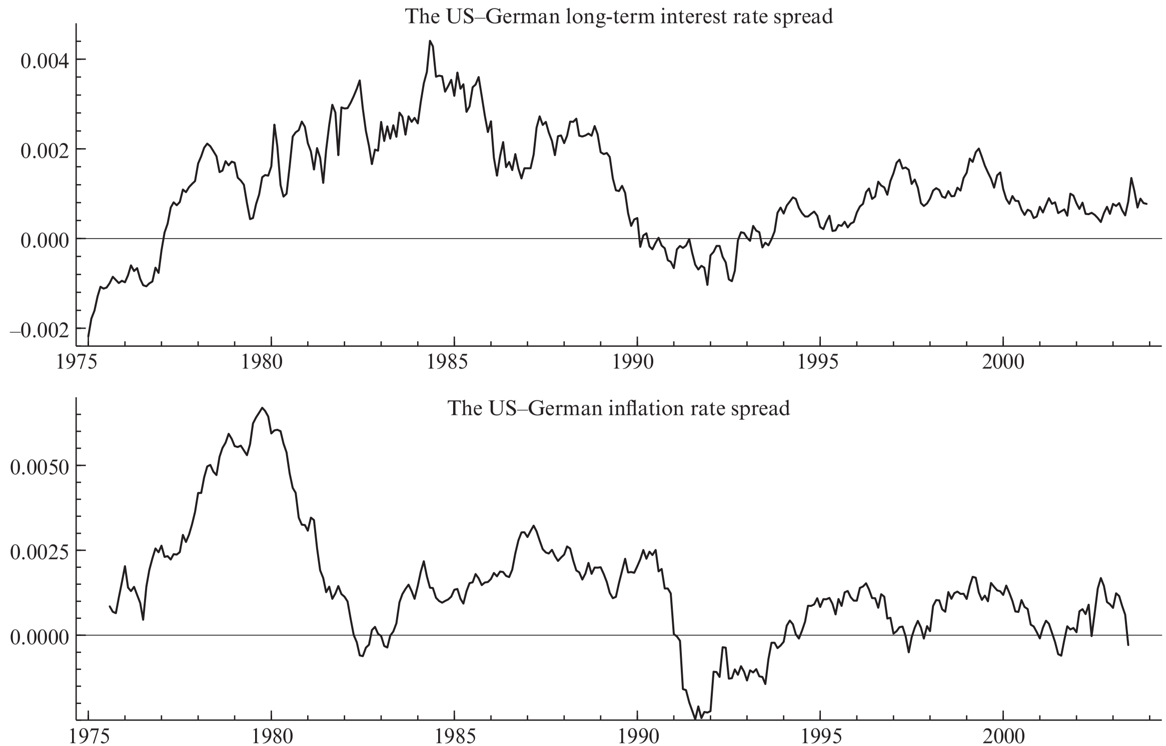

Figure 17.4 *US–German long-term interest rate differential (upper panel) and US–German inflation rate differential (lower panel)*

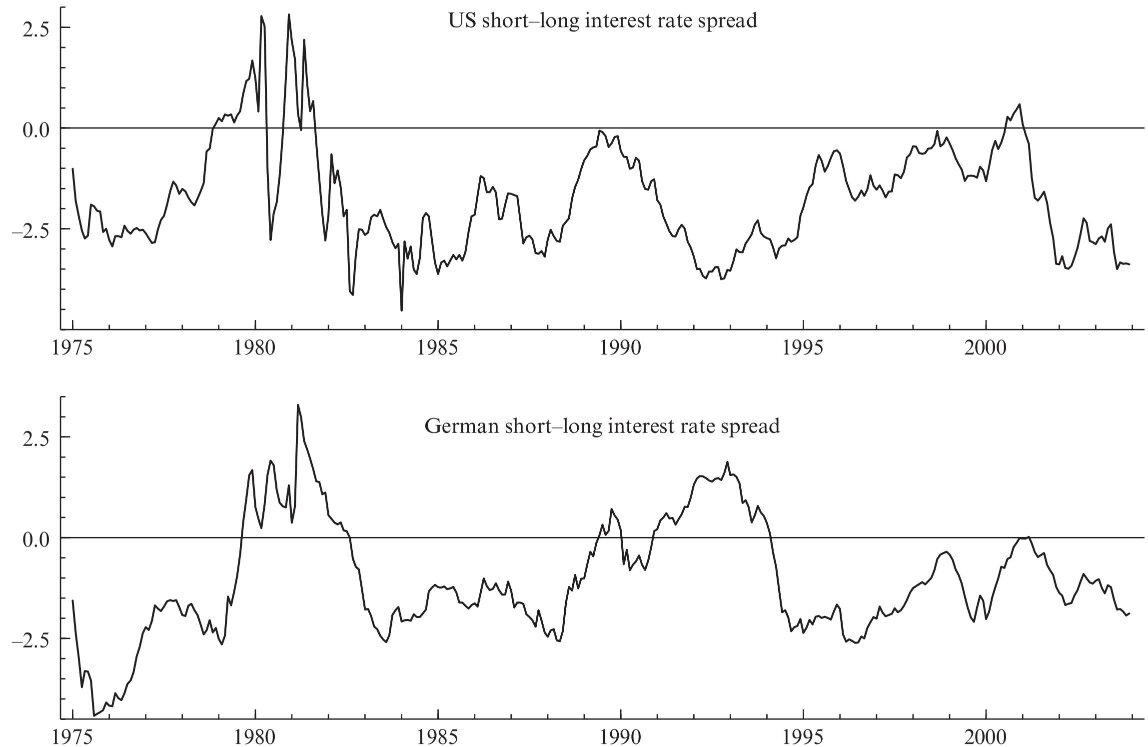

Figure 17.5 *The US short–long interest rate spread (upper panel) and the German short–long interest rate spread (lower panel)*

spreads exhibit a similar persistence. Based on a CVAR analysis of five US zero coupon bonds, Giese (2008) finds that the term spreads are non-stationary but cointegrated, and that shocks to the level and the slope of the term structure are pushing the system.

To dig deeper into these data regularities, Johansen and Juselius (1992) and Juselius (1995) first tested a number of hypotheses related to the PPP and uncovered interestrate parity (UIP) based on US–UK data. This was followed up by a similar study of US–Japanese data (Juselius and MacDonald, 2004) and one of US–German data (Juselius and MacDonald, 2007). All of them find that the persistent movements in the real exchange rate are compensated by similar persistent movements in the interest rate differential (either short term or long term).

How can this be empirically and theoretically understood? Obviously, prices and nominal exchange rates differ in one important respect: nominal exchange rates are primarily being determined in the speculative market for foreign exchange based on future expectations of capital gains, whereas prices of tradables are determined in an internationally highly competitive goods market. That prices and nominal exchange rates behave differently has, of course, also been recognized by many REH-based models such as the Dornbush and Frankel overshooting type of models (Dornbusch, 1976; Dornbusch and Frankel, 1995) some of which allow for persistence in the real exchange rate (Benigno, 2004; Engel and West, 2006; Galí and Monacelli, 2005). But, Juselius (2009b) finds that the swings in the real exchange rate are too persistent for the REH-based models to tell an empirically credible story. The long swings remain a puzzle in these models. In contrast, the IKE-based models proposed by Frydman and Goldberg (2007, 2008) seem to provide a remarkably good fit. See Frydman et al. (2008) and Juselius (2009b).

To test the two competing theories Juselius (2010) formulated a theory-consistent CVAR scenario for REH-based and IKE-based monetary models. The scenario analysis shows that under REH the real exchange rate and the real interest rate differential are each I(0) or at most near I(1), whereas under IKE they are I(1) or near I(2) and cointegrate to stationarity. One important testable difference is that the degree of persistence is one degree higher under IKE than under REH. Strong evidence for the IKE-based model is found in Frydman et al. (2008), Johansen et al. (2010), and Juselius (2010).

In the stock market, the efficient market hypothesis is one of the theoretical cornerstones of financial behavior. It has often been formulated as a random walk hypothesis of stock prices, which has been tested and often not rejected. The random walk hypothesis, however, is much too simplistic as a description of financial market behavior and as a test of the EMH. A more sophisticated formulation was given in the seminal article by Campbell and Shiller (1987) relying on a REH present value model for stock prices and nominal exchange rates. The model is based on the assumption that both prices and dividends are first-order difference stationary and that parameters are constant over the full sample period.[13]

In a still ongoing project, the author of this chapter applied the CVAR model to the Shiller data and obtained the following tentative results: (1) a constant equilibrium relationship over the entire sample 1934–2008 is strongly rejected, but reasonably constant subsamples can be found; (2) first-order difference stationarity is rejected in favor of second-order difference stationarity; (3) dividends alone cannot explain movements in stock prices, earnings are also needed; (4) stock prices tend to fluctuate in a boom and bust fashion around the price–earnings ratio; (5) earnings are influenced by the price–dividend ratio and are error-correcting toward the price–earnings ratio. Thus, rising earnings lead to rising stock prices, and rising stock prices has a positive 'long-run' effect on earnings which, therefore, feed back on stock prices.

These empirical regularities, though still tentative, resemble many of those found in the foreign currency markets and suggests that an IKE-based model might also work well in the stock market. For example, under IKE, we would expect that financial behavior is non-constant over time, that endogenously loss-averse agents produce near I(2) persistent swings in asset prices, and that this would likely imply strong feedback effects between prices, dividends and earnings. This is what data show.

17.6.3 Financial Markets and Macroeconomic Modeling

Speculative behavior in the foreign exchange market (through its impact on the real exchange rate, the real interest rate and the term spread) is bound to have strong implications for macroeconomic modelling and policy. This is also the story macroeconomic that data tell. The non-stationarity of the term spread and the real interest rate (and indirectly the real exchange rate) is likely to have a strong impact on monetary policy transmission mechanism. Empirical results in Juselius (1992, 1996, 1998a, 1998b, 1999, 2001, 2004, 2006), Juselius and Toro (2005) and Juselius and Johansen (2006) demonstrate that standard transmission mechanisms for how changes in monetary instruments transmit through the system are seriously impaired by the interference of financial market behavior.

The non-stationarity of the real exchange rate and the real interest rate have been shown to have a strong impact on domestic wage and price setting[14] with important implications for fiscal policy. For example, Juselius (2006, 2011b) and Juselius and Ordóñez (2009) provide strong empirical evidence. Further evidence supporting the importance of these findings for policy can be found in Bjørnstad and Nymoen (2008), Fanelli (2008) and M. Juselius (2008), showing essentially no empirical evidence in support of the new Keynesian Phillips curve model favored by central banks all over the world.

The effect of stock prices on the real economy is also potentially large through their impact on earnings. The latter is likely to have a strong impact on financial and housing wealth and, thus, on consumption and investment behavior. The results in Tabor (2009) suggest that this may explain why the domestic variables of the Ireland model were not able to explain the data variation over the more recent period.

Thus, the stories that data tell when allowed to speak freely seem to reject REH-based models, to be consistent with an IKE based framework, and to suggest significant interaction between the financial and real sectors of the economy. I am convinced that a Sherlock Holmesian type of approach has the ability to inspire new economic thinking in empirically relevant directions.

17.7 CAN WE TRUST THE STORIES DATA TELL?

Claiming that one can learn anything from data analysis is often considered controversial in economics. Many economists would argue that unless the empirical model is constrained by theory from the outset, one would not be able to make sense of the results: without the mathematical logic of the theory model, one opens up possibilities of quackery. For example, as some of the CVAR results reported in Juselius and Franchi (2007) were less straightforward to interpret, one might ask, as one of the referees did, whether at the end of the day DSGE modeling with interpretable theory is to be preferred to VAR modeling without theory. No doubt many economists would be sympathetic to such a view, and sometimes for good reasons: there is an abundance of bad VAR applications in the literature. These give the impression of having been applied mechanically (pressing the VAR button) rather than asking sharp and precise questions about the economy. Ireland's VAR(1) process was just an add-on to the RBC model without any attempt to use it as a source of valuable data information.

Such VAR modeling has nothing to do with a maximum-likelihood-based VAR analysis. To claim that the statistical analysis is based on full information maximum likelihood requires that the model satisfactorily describes all aspects of the data. This means that the researcher must carefully check for many things: have there been shifts in mean growth rates or in equilibrium means? Are the effects of interventions, reforms, and changing policy properly modelled? Is the sample period defining a constant parameter regime? Is the information set correctly chosen? and so on. The accuracy of the results depends on all this being correct in the model. To make the necessary analysis to develop a satisfactory model is a time-consuming and tedious process that depends upon the researcher's judgement and expertise and has nothing to do with pressing the VAR button. But without such checking, the results can be (and often are) close to useless and if they are taken seriously by policymakers, even worse than worthless.

It is, therefore, important to emphasize that a statistically adequate VAR analysis has to obey equally strict scientific rules as an analysis of a mathematical model in economics. In principle there is no arbitrariness in such empirical analyses, as Spanos (2009) points out. However, objectivity can only be achieved provided data are not constrained from the outset in theoretically prespecified directions as it would otherwise be impossible to know which results are due to the assumptions made and which are the empirical facts. The only way the methodology works properly is by allowing the data to speak as freely as possible about empirical regularities. This is, of course, not the same as letting the data speak without any theory, which generally would not produce anything useful.

Another frequent argument is that the quality of economic data is too low. I agree that economic time series data seldom correspond exactly to the theoretical concepts of a theory model. For example, the representative agent's income, consumption and hours worked in Ireland's model has little in common with the various measurements of aggregate income, private consumption and total hours worked that are used in his empirical analysis. While, admittedly, macro data are contaminated with measurement errors, such errors may not be of great concern for the more important long-run analysis, unless they are systematic and cumulate to non-stationarity. Whatever the case, theoretically correct measurements do not exist and, hence, cannot be used by politicians and decision makers to react on.

17.8 THE ROLE OF EVIDENCE AND THEORY IN MACROECONOMICS: A CONCLUDING DISCUSSION

For more than a decade, the DSGE modelling has been the preferred way of doing empirical macroeconomics. But although the models have become a lot more sophisticated over time, they were not able to warn policymakers about the approaching crisis in 2007, the worst in almost a century.

This failure revealed almost painfully that extant economic models are seriously lacking when it comes to understand out-of-equilibrium behavior. As the Dahlem report (Colander et al., 2009) expressed it: 'When the crisis struck policy makers were essentially left to grope in the dark without much guidance for how to ride out the crises. They could only hope that their costly policy measure would have the intended effect.'

Why did the models perform so badly, exactly when reliable advice was most urgently needed? I have argued that one reason these models fail to describe macroeconomic behavior adequately is that they are generally based on information rich agents with 'rational expectations' and, therefore, are consistent with fast adjustment back towards equilibrium when pushed away by exogenous shocks. One could say that if the REH hypothesis was a good description of how economic agents behave, data would not have exhibited the persistent behavior away from long-run steady states that preceded the recent crisis. The fact that the empirical analysis of Ireland's model at a first glance seemed impressive, despite its strong empirical rejection, is therefore a warning against drawing conclusions from models based on many untested assumptions. Strong economic priors imposed on the data without testing may say more about the beliefs of the researcher than the state of the economic reality. The ultimate question is, therefore, whether we can afford to let economic policy be guided by beliefs which are not strongly backed up by empirical evidence.

But can we do better? The fact that economic data are often well described by CVAR models may suggest that empirically relevant economic models need to be formulated as dynamic adjustment models in growth rates and equilibrium errors, the so-called equilibrium correction models. See, for example, Hendry (1987, 1995) and Juselius (2006). Such models are inherently consistent with the idea that unanticipated shocks cumulate over time to generate stochastic trends in the variables. However, the fact that basic long-run relationships such as purchasing power parity, real interest rates, uncovered interest rate parity, the term spread, the consumption–income ratio, the capital–income ratio, the labour–income ratio, the natural rate of the Phillips curve, and so on. often exhibit a pronounced persistence untenable with stationarity, suggests that empirically relevant models need to be formulated in acceleration rates, medium-run relations between growth rates, and dynamic rather than static long-run relations. In the econometrics jargon it means that we should formulate models for an I(2) rather than an I(1) world.

The economic rationale for choosing such models is that an imbalance (a disequilibrium error) in one sector of the economy cannot develop into a persistent movement unless compensated by something else. For example, the persistent out-of-equilibrium behavior of the real exchange rate was possible because of a compensating movement in the relative real interest rates. On the other hand, a persistent movement in the real interest rate is likely to generate persistent movements elsewhere, for example in the unemployment rate. Therefore, to understand the pronounced persistence of economic time-series we need to unwind the mechanisms that have generated it. By its ability to uncover empirically stable combinations of individually persistent (equilibrium) errors, cointegration analysis is potentially a very efficient procedure in this respect. It suggests where to start digging.

But the persistence of static equilibrium errors have implications for economic modelling that take us even further. In such disequilibrium situations, the mathematical logic of the CVAR model tells us that the adjustment back towards long-run sustainable equilibrium states will entail growth rates that move in a similarly persistent manner (Kongsted, 2005; Juselius, 2006). The implication is that we need to move from a framework of static to dynamic equilibrium relationships, in particular, when the horizon of the analysis is the medium run of say five or six years. The fact that we find static equilibrium errors to be unstable (non-stationary) does not preclude the existence of stable relationships; it just moves the discussion of stability to a higher level. In this sense, a standard static long-run relation can be seen as a special case of a more general dynamic equilibrium relation.

Can a DSGE model be empirically assessed? Many of these models would also allow for permanent shocks that cumulate to stochastic trends, for example shocks to technology and preferences. But, in this case, the non-stationarity of the data is incorporated in the model by assuming an exogenously given stochastic trend. A crucial difference between the two approaches is that the number of stochastic trends is estimated in the CVAR model, not assumed, and the presumption that one of them is a technology trend would be formulated as a testable restriction on the parameters of the CVAR model rather than imposed from the outset. But, because the general structure of a DSGE and a CVAR model is similar, the former can in principle be formulated as a submodel within the CVAR and is, therefore, testable as illustrated for the DSGE model in Ireland (2004a). That most assumptions underlying the model were rejected and the conclusions reversed seem to illustrate Tony Lawson's claim that 'models that get "the right results" and "address interesting questions" may nevertheless turn out to be misleading and empirically irrelevant' (Lawson, 2009).

I have argued here that one reason the DSGE models did so poorly in forecasting the financial and economic crisis of 2007–09 was because they ignored the importance of interactions between the financial and the real sectors of the economy. Would CVAR modeling have done better? The crisis was a consequence of equilibrium errors that were allowed to cumulate over a fairly long period without a proper warning system which, at an early stage, would have signaled that the system was moving seriously out of equilibrium. Already many years before the bubble burst, CVAR analyses showed that the relative house–consumption price relationship exhibited very pronounced non-stationary behavior and that this was (primarily) facilitated by increasing financial wealth. Prior to the bursting of the bubble, the house–consumption price ratio increased almost exponentially, signalling that house prices were moving far away from their sustainable level as given by the very low levels of Consumer Prices Index (CPI) inflation and interest rates. This would have been absolutely the last moment for the authorities to react to prevent the subsequent disaster. But the fact that the extraordinarily high level of house prices was sustainable only by a record high credit expansion, and extraordinarily low levels of nominal interest rates and inflation rates, should have been a reason for concern much earlier.

My point here is that in order to prevent the next crisis, we need a proper understanding of the disequilibrium forces at the macro level. Markets are infinitely inventive and it is almost impossible to foresee what could generate the next crisis. Unless we build up a good macroeconomic understanding and a reliable signalling system, it is very likely that the next crisis will come as an equally big surprise. One way of avoiding this happening is to learn from the data in a systematic and structured way using available theories and hypotheses, but at the same time being open to signals in data suggesting that there are mechanisms we do not yet understand. By embedding the theory model in a broader empirical framework, the empirical analysis should be able to point to possible pitfalls in macroeconomic reasoning, and at the same time to generate new hypotheses for how to modify too narrowly specified theoretical models. The clues to present and previous crises are hidden in the historical data and we need to take them much more seriously than is usually done.

NOTES

1. Useful comments by Michael Goldberg, Søren Johansen and Mikael Juselius are gratefully acknowledged.
2. Because the technology shock is the only source of randomness in the model, the model is stochastically singular.
3. The simulation in Johansen (2006) shows that with a root of 0.9987, at least 5000 observations are needed to get close to the correct size of a test on the steady-state value.
4. These were not reported in Ireland, but derived from his model.

5. Gross capital formation is used as a measure of capital instead of the Ireland variable which was generated to fit the RBC hypothesis.
6. Transitory shocks which do not cumulate are called I(0), permanent shocks which cumulate once are called I(1), and permanent shocks that cumulate twice are called I(2).
7. A similar transformation of the theoretical model into growth rates and cointegration is suggested in Ireland (2004b).
8. The second period contained a lot more persistence, and as will be illustrated below it is questionable whether there is any cointegration.
9. To cite an anonymous discussant to Spanos (2009): 'I don't think DSGE modelers don't care about the empirical support of their theories; they just have a different metric to assess the usefulness of their theories. If the standard metric rejects the model, just choose a different metric.'
10. This points to the importance of not just assuming that the model contains all basic relevant variables, but checking whether this is the case, for example by using automatic search procedures such as autometrics (Hendry and Krolzig, 2005).
11. The fact that the literature abounds with econometric tests suggesting that the unit root hypothesis can be rejected does not make the pronounced persistence in the data disappear. It may, however, say something about the way econometrics is used to illustrate prior beliefs, rather than as a tool for discovery.
12. These are known as the PPP puzzle, the long swings puzzle and the nominal exchange rate disconnect puzzle (Rogoff, 1996).
13. Shiller (2000), in a critique of EMH, concluded that financial markets are essentially driven by irrational exuberance. But under this hypothesis, we would essentially have to give up understanding financial market behavior and how to best cope with it.
14. This provides evidence for some important hypotheses in Phelps (1994).

REFERENCES

Andersen, T.W. (1991), 'Tryggve Haavelmo and Simultaneous Equation Models', *Scandinavian Journal of Statistics*, 18, 1–19.

Backhouse, R.E. and A. Salanti (2000a), 'Macroeconomics and the Real World: Volume 1: Econometric Techniques and Macroeconomics'. Oxford: Oxford University Press.

Backhouse, R.E. and A. Salanti (2000b), 'Macroeconomics and the Real World: Volume 2: Keynesian Economics, Unemployment, and Policy'. Oxford: Oxford University Press.

Benigno, G. (2004), 'Real Exchange Rate Persistence and Monetary Policy Rules'**,** *Journal of Monetary Economics*, **51**, 473–502.

Bjørnstad, R. and R. Nymoen (2008), 'The New Keynesian Phillips Curve Tested on OECD Panel Data', *Economics: The Open-Access, Open-Assessment E-Journal*, **2(**23), http://www.economics-ejournal.org/economics/journalarticles/2008-23..

Campbell, J.Y. and R. Shiller (1987), 'Cointegration and Tests of Present Valus models', *Journal of Political Economy*, **95**(5), 1062–1089.

Clements, M.P. and D.F. Hendry (1999), *Forecasting Non-Stationary Time Series*, Cambridge, MA: MIT Press.

Clements, M.P. and D.F. Hendry (2008), 'Economic Forecasting in a Changing World', *Capitalism and Society*, **3**, 1–18.

Cochrane, J. (2009), 'How Did Paul Krugman Get It so Wrong?', http://modeledbehavior.com/2009/09/11/john-cochrane-responds-to-paul-krugman-full-text/.

Colander, D. (ed.) (1996), *Beyond Microfoundations: Post Walrasian Macroeconomics*, Cambridge: Cambridge University Press.

Colander, D. (ed.) (2000), *The Complexity Vision and the Teaching of Economics* Cambridge: Cambridge University Press.

Colander, D. (2001), *'The Lost Art of Economics. Essays on Economics and the Economics Profession'*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Colander, D. (ed.) (2006), *Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*, Cambridge: Cambridge University Press.
Colander, D., M. Goldberg, A. Haas, K. Juselius, A. Kirman, T. Lux, B. Sloth (2009), 'The Financial Crisis and the Systemic Failure of the Economics Profession', *Critical Review*, **21**(3), 249–67.
Colander, D., P. Howitt, A. Kirman, A. Leijonhufvud and P. Mehrling (2008), 'Beyond DSGE Models: Toward an Empirically Based Macroeconomics' , *American Economic Review*, **98**, 2.
Cooley, T.F. and S.F. Leroy (1985), 'Atheoretical Macroeconometrics: A Critique', *Journal of Monetary Economics*, **16**(3), 283–308.
Dornbusch, R. (1976), 'Expectations and Exchange Rate Dynamics', *Journal of Political Economy*, **84**(7), 1161–1174.
Dornbusch, R. and J.A. Frankel (1995), 'The Flexible Exchange Rate System: Experience and Alternatives', in S. Borner (ed.), *International Finance and Trade*, London: Macmillan, reprinted in J.A. Frankel (ed.), *On Exchange Rates*, Cambridge, MA: MIT Press.
Engel, C. and K. West (2006), 'Taylor Rules and the Deutschemark-Dollar Real Exchange Rate', *Journal of Money, Credit and Banking*, **38**, 1175–1194.
Engle, R.F. and C.W.J. Granger (1987), 'Cointegration and Error Correction: Representation, Estimation, and Testing', . *Econometrica*, **55**(2), 251–276.
Fanelli, L. (2008), 'Evaluating the New Keynesian Phillips Curve under VAR-based Learning', *Economics: The Open-Access, Open-Assessment E-Journal*, **2**(33), http://www.economics-ejournal.org/economics/journalarticles/2008-33.
Frydman, R. and M. Goldberg (2007), *Imperfect Knowledge Economics: Exchange rates and Risk*, Princeton, NJ: Princeton University Press.
Frydman, R. and M. Goldberg (2008), 'Macroeconomic Theory for a World of Imperfect Knowledge', *Capitalism and Society*, **3**(1), http://bepress.com/cas/vol3/iss3/art1.
Frydman, R. and M. Goldberg (2009), 'Financial Markets and the State: Price Swings, Risk, and the Scope of Regulation', *Capitalism and Society*, **4**(2), http://bepress.com/cas/vol4/iss2/art2.
Frydman, R. and M. Goldberg (2011), *Beyond Mechanical Markets: Risk and the Role of Asset Price Swings*, Princeton, NJ: Princeton University Press.
Frydman, R., M. Goldberg, K. Juselius, and S. Johansen (2008), 'Imperfect Knowledge and a Resolution of the Purchasing Power Parity Puzzle: Empirical Evidence', Working Paper 08-31, University of Copenhagen.
Galí, J. and T. Monacelli (2005), 'Monetary Policy and Exchange Rate Volatility in a Small Open Economy', *Review of Economic Studies*, **72**, 707–734.
Giese, J.V. (2008), 'Level, Slope, Curvature: Characterising the Yield Curve in a Cointegrated VAR Model', *Economics: The Open-Access, Open-Assessment E-Journal*, **2**(28), http://www.economics-ejournal.org/economics/journalarticles/2008-28.
Gilbert, C.L. (1986), 'Professor Hendry's Econometric Methodology', *Oxford Bulletin of Economics and Statistics*, **48**, 283–307.
Haavelmo, T. (1944), 'The Probability Approach to Econometrics', *Econometrica*, **12**(Supplement), 1–118.
Hendry, D.F. (1987), 'Econometric Methodology: A Personal Perspective', in T.F. Bewley (ed.), *Advances in Econometrics*, Cambridge, MA: Cambridge University Press, Chapter 10.
Hendry, D.F. (1995), *Dynamic Econometrics*, Oxford: Oxford University Press.
Hendry, D.F. (2009), 'The Methodology of Empirical Econometric Modeling: Applied Econometrics Through the Looking-Glass', in T.C. Mills and K. Patterson (eds), *Palgrave Handbook of Econometrics, Vol. 2*, London: Macmillan, pp. 3–67.
Hendry, D.F. and K. Juselius (1999), 'Explaining Cointegration Analysis. Part 1', *Energy Journal*, **21**(1), 1–42.

Hendry, D.F. and K. Juselius (2000), 'Explaining Cointegration Analysis. Part 2', *Energy Journal*, **22**(1), 1–52.
Hendry, D.F. and H-M. Krolzig (2005), 'The Properties of Automatic Gets Modelling', *Economic Journal*, **115**, C32–C61.
Hendry, D.F. and G.E. Mizon (1993), 'Evaluating Econometric Models by Encompassing the VAR', in P.C. Phillips (ed.), *Models, Methods and Applications of Econometrics*, Oxford: Blackwell, pp. 272–300.
Hendry, D.F. and G.E. Mizon (2011), 'What Needs Rethinking in Macroeconomics', *Global policy*, **2**(2), 176–183.
Hendry, D.F., A. Spanos and N.R. Ericsson (1989), 'The Contributions to Econometrics in Tryggve Haavelmo's "The Probability Approach in Econometrics"', *Socialøkonomen*, **43**(11), 12–17.
Hoover, K. (2006), 'The Past as Future: The Marshallian Approach to Post Walrasian Econometrics', in D. Colander (ed.), *Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*, Cambridge: Cambridge University Press, pp. 239–257.
Hoover, K., S. Johansen and K. Juselius (2008), 'Allowing the Data to Speak Freely: The Macroeconometrics of the Cointegrated VAR', *American Economic Review*, **98**, pp. 251–255.
Ireland, P.N. (2004a), 'A Method for Taking Models to the Data', *Journal of Economic Dynamics and Control*, **28**(6), 1205–1226.
Ireland, P.N. (2004b), 'Technology Shocks in the New Keynesian Model'**,** *Review of Economics and Statistics*, **84**(4), 923–936.
Johansen, S. (1988). 'Statistical Analysis of Cointegration Vectors', *Journal of Economic Dynamics and Control*, **12**(213), 231–254.
Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Oxford University Press.
Johansen, Søren (2006), 'Confronting the Economic Model with the Data', in D. Colander (ed.), *Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*, Cambridge: Cambridge University Press, pp. 287–300.
Johansen, S. and K. Juselius (1992), 'Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and the UIP for UK', *Journal of Econometrics*, **53**, 211–244.
Johansen, S., K. Juselius, R. Frydman and M. Goldberg (2010), 'Testing Hypotheses in an I(2) Model With Piecewise Linear Trends. An Analysis of the Persistent Long Swings in the Dmk/$ Rate', *Journal of Econometrics*, **158**, 117–129.
Juselius, K. (1992), 'Domestic and Foreign Effects on Prices in an Open Economy. The Case of Denmark', *Journal of Economic Policy Modelling*, **14**, 401–28; reprinted (1995) in Ericsson and J.S. Irons (eds), *Testing Exogeneity*. Advanced Texts in Econometrics, Oxford: Oxford University Press, pp. 161–90.
Juselius, K. (1995), 'Do Purchasing Power Parity and Uncovered Interest Rate Parity Hold in the Long Run? An Example of likelihood Inference in a Multivariate Time Series Model', *Journal of Econometrics*, **69**(1), 211–40.
Juselius, K. (1996), 'An Empirical Analysis of the Changing role of German Bundesbank after 1983', *Oxford Bulletin of Economics and Statistics*, **58**, 791–817.
Juselius, K. (1998a), 'Changing Monetary Transmission Mechanisms Within the EU', *Empirical Economics*, **23**, 455–481.
Juselius, K. (1998b) 'A Structured VAR under Changing Monetary Policy', *Journal of Business and Economics Statistics*, **6**, 400–412.
Juselius, K. (1999), 'Models and Relations in Economics and Econometrics', *Journal of Economic Methodology*, **6**(2), 259–290.
Juselius, K. (2001) 'European Integration and Monetary Transmission Mechanisms: The Case of Italy', *Journal of Applied Econometrics*, **16**, 341–358.
Juselius, K. (2004), 'Inflation, Money Growth, and I(2) Analysis', in A. Welfe (ed.) *Contributions to Economic Analysis. New Directions in Macromodelling: Essays in Honor of J. Michael Finer*, Bingley: Emerald, pp. 69–106.

Juselius, K. (2006), *The Cointegrated VAR Model: Methodology and Applications*, Oxford: Oxford University Press.
Juselius, K. (2009a), 'Special Issue on Using Econometrics for Assessing Economic Models – An Introduction', *Economics: The Open-Access, Open-Assessment E-Journal*, **3**(2009–28).
Juselius, K. (2009b), 'The Long Swings Puzzle. What the Data Tell when Allowed to Speak Freely', in T.C. Mills and K. Patterson (eds), *The New Palgrave Handbook of Econometrics, Vol. 2*, London: Macmillan, pp. 349–84.
Juselius, K. (2010), 'Testing Exchange Rate Models Based on Rational Expectations versus Imperfect Knowledge Economics: A Scenario Analysis', Working Paper, Economics Department, University of Copenhagen.
Juselius, K. (2011a), 'Time to Reject the Privileging of Economic Theory over Empirical Evidence? A Reply to Lawson', *Cambridge Journal of Economics*, **35**(2), 423–436.
Juselius, K. (2011b), 'Imperfect Knowledge, Asset Price Swings and Structural Slumps: A Cointegrated VAR Analysis of Their Interdependence', Department of Economics, University of Copenhagen, Discussion papers 10-31; forthcoming in Phelps, E. and Frydman, R. (eds), *Microfoundations for Macroeconomic Theory*, Princeton University Press.
Juselius, K. and M. Franchi (2007), 'Taking a DSGE Model to the Data Meaningfully', *Economics-The Open-Access, Open-Assessment E-Journal*, 2007-4, http://www.economics-ejournal.org/economics/journalarticles/3007-4.
Juselius, K. and Johansen, S. (2006), 'Extracting Information from the Data: A European View on Empirical Macro', in D. Colander (ed.), *Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*, Cambridge: Cambridge University Press, pp. 301–334.
Juselius, K. and R. MacDonald (2004), 'The International Parities Between USA and Japan', *Japan and the World Economy*, **16**, 17–34.
Juselius, K. and R. MacDonald (2007), 'International Parity Relationships Between Germany and the United States: A Joint Modelling Approach', in A. Morales-Zumaquero (ed.), *International Macroeconomics: Recent Developments*. New York: Nova Science Publishers, pp. 79–104.
Juselius, K. and J. Ordóñez (2009), 'Balassa-Samuelson and Wage, Price and Unemployment Dynamics in the Spanish Transition to EMU Membership', *Economics: The Open-Access, Open-Assessment E-Journal*, **3**(4), http:/www.economics-ejournal.org/economics/journalarticles/2009-4.
Juselius, K. and J. Toro (2005), 'The effect of joining the EMS. Monetary transmission mechanisms in Spain', *Journal of International Money and Finance*, **24**, 509–531.
Juselius, M. (2008), 'Testing the New Keynesian Model on US and Euro Area Data', *Economics: The Open-Access, Open-Assessment E-Journal*, **2**(24), http://www.economics-ejournal.org/economics/journalarticles/2008-24.
Katlesky, A. (2009), 'Goodbye, Homo Economicus', *Prospect Magazine*, **157**, 26 April.
Keynes, J.M. (1937), 'General Theory of Employment', *Quarterly Journal of Economics*, **51**(2), 209–223.
Kongsted, H.C. (2005), 'Testing the Nominal-to-Real Transformation', *Journal of Econometrics*, **124**, 205–225.
Krugman, P. (2009), 'How did Economists Get It so Wrong?', Column in *New York Times*, 6 September, http://www.nytimes.com/2009/09/06/magazine/06Economic-t.html?_r=4.
Lawson, T. (2009), 'The Current Economic Crisis: Its Nature and the Course of Academic Economics', *Cambridge Journal of Economics*, **33**, 759–777.
Leijonhufvud, A (2009), 'Out of the Corridor: Keynes and the Crisis', *Cambridge Journal of Economics*, **33**(4), 741–757.
Møller, N.F. (2008), 'Bridging Economic Theory Models and the Cointegrated Vector Autoregressive Model', *Economics: The Open-Access, Open-Assessment E-Journal*, **2(**36), http://www.economics-ejournal.org/economics/journalarticles/2008-36.
Pagan, A.R. (1987), 'Three Econometric Methodologies: A Critical Appraisal', *Journal of Economic Surveys*, **1**, 3–24.

Phelps, E. (1994), *Structural Slumps*, Princeton, NJ: Princeton University Press.
Rogoff, K. (1996), 'The Purchasing Power Parity Puzzle', *Journal of Economic Literature*, **34**, 647–668.
Shiller, R.J. (2000), *Irrational Exuberance.* New York: Broadway Books.
Soros, G. (1987), *The Alchemy of Finance*, New York: John Wiley & Sons.
Spanos, A. (1995), 'On Theory Testing in Econometrics: Modeling with Nonexperimental Data', *Journal of Econometrics*, **67**, 189–226.
Spanos, A. (2006), 'Econometrics in Retrospect and Prospect', in T.C. Mills and K. Patterson (eds), *Palgrave Handbook of Econometrics*, Vol. 1, London: MacMillan, pp. 3–58.
Spanos, A. (2009), 'The Pre-Eminence of Theory versus the European CVAR Perspective in Macroeconometric Modeling', *Economics: The Open-Access, Open-Assessment E-Journal*, **3**(10), http://www.economics-ejournal.org/economics/journalarticles/2009-10.
Summers, L. (1991), 'The Scientific Illusion in Empirical Macroeconomics', *Scandinavian Journal of Economics,* **93**, 129–148.
Tabor, Nyboe M. (2009), 'The Effect of Financial Markets on US Consumption: A Cointegrated VAR Approach. 1954–2008', Masters thesis, Department of Economics, University of Copenhagen.
Yule, U. (1926), 'Why do we Sometimes Get Nonsense-Correlations Between Time Series? A Study in Sampling and the Nature of Time Series', *Journal of the Royal Statistical Society*, **89**, 1–63.