Observing Shocks[1]

Pedro Garcia Duarte† Kevin D. Hoover‡

RESUMO:

Choque é um termo corriqueiro em economia e aparece em cerca de vinte e cinco por cento de todos os artigos em economia, e cerca de metade dos artigos de macroeconomia. O significado da palavra é diverso e usado das várias maneiras ao longo de todo o período. Surpreendente, no entanto, é que o aumento substancial do uso da palavra "choque" ocorreu no início da década de 1970. Este artigo apresenta a história do uso de choques em macroeconomia, de Frisch e Slutzky nas décadas de 1920 e 1930 até o novo consenso em macroeconomia (os modelos DSGE) onde choques geram funções de resposta a impulso que são usadas para estimar os parâmetros dos modelos, passando pelos teóricos dos ciclos econômicos reais. Tal história se organiza em torno da "observabilidade" dos choques e mostra desenvolvimentos conceituais críticos em economia. Ela também serve como um estudo de caso que ilustra e questiona a distinção entre "dados" e "fenômenos", feita pelos filósofos da ciência James Bogen e James Woodward. A história dos choques mostra que esta distinção deve ser substancialmente relativizada, para poder ser aplicada plausivelmente em economia.

Palavras-Chave: choques, macroeconomia novo clássica, modelos DSGE, funções de resposta a impulso

ABSTRACT:

Shock is a term of art that pervades modern economics appearing in nearly a quarter of all journal articles in economics and in nearly half in macroeconomics. Surprisingly, its rise as an essential element in the vocabulary of economists can be dated only to the early 1970s. The paper traces the history of shocks in macroeconomics from Frisch and Slutzky in the 1920s and 1930s through real-business-cycle and DSGE models and to the use of shocks as generators of impulse-response functions, which are in turn used as data in matching estimators. The history is organized around the observability of shocks. As well as documenting a critical conceptual development in economics, the history of shocks provides a case study that illustrates, but also suggests the limitations of, the distinction drawn by the philosophers of science James Bogen and James Woodward between data and phenomena. The history of shocks shows that this distinction must be substantially relativized if it is to be at all plausible.

Keywords: shock, new classical macroeconomics, DSGE model, impulse-response function

Área ANPEC: 01 **Classificação JEL:** B22, B23, B41

† Department of Economics, University of São Paulo (USP), Av. Prof. Luciano Gualberto 908, Cidade Universitária, São Paulo, SP, 05508-010, Brazil. E-mail: pgduarte@usp.br; Tel. +55 (11) 3091-5944. I gratefully acknowledge financial support from FAPESP and CNPq (Brazil).

‡ Department of Economics and Department of Philosophy, Duke University, Box 90097, Durham, NC 27708-0097, U.S.A. E-mail: kd.hoover@duke.edu; Tel. +(919) 660-1976. I acknowledge the support of the U.S. National Science Foundation (grant no. NSF SES-1026983).

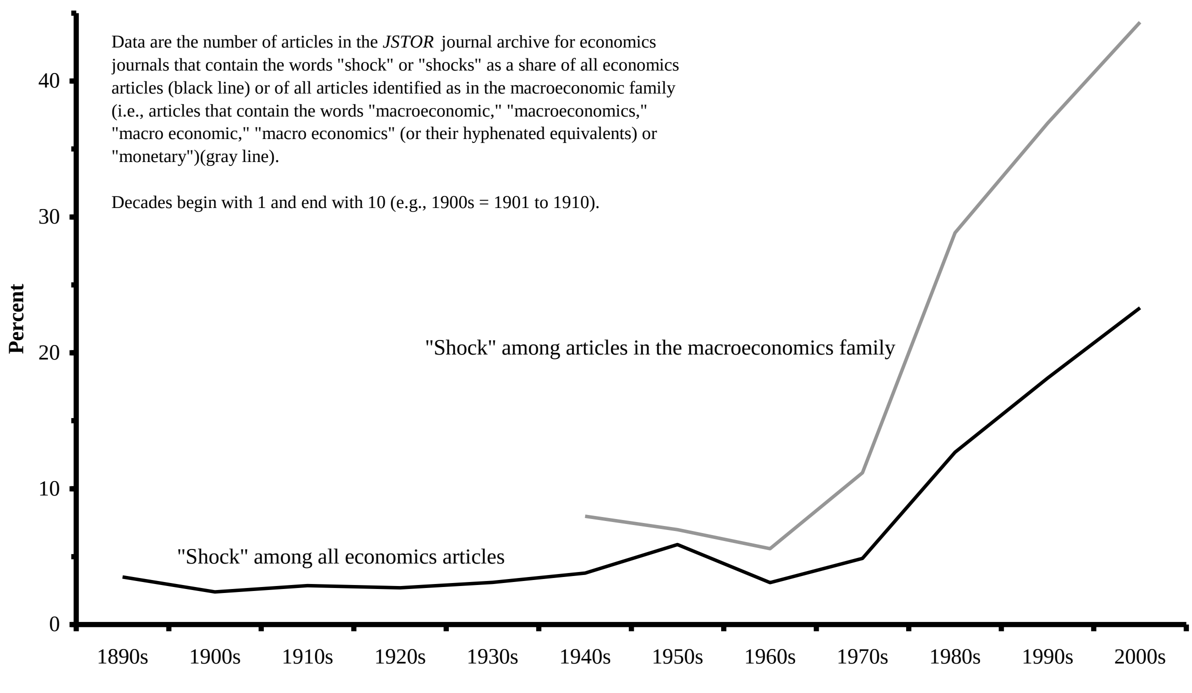
[1] A longer version of this paper is available on SSRN (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1840705).

Observing Shocks[1]

I. The Rise of Shocks

Shock is a relatively common English word – used by economists for a long time and to a large extent much as other people used it. Over the past forty years or so, economists have broken ranks with ordinary language and both narrowed their preferred sense of *shock* and promoted it to a term of econometric art. The black line in Figure 1 shows the fraction of all articles using the term "shock" (or "shocks") as a percentage of all articles in the economics journals archived in the *JSTOR* database from the last decade of the 19th century to the present. The striking feature of the figure is that the use of "shock" varies between 2½ percent and 5 percent up to the 1960s, and then accelerates steeply, so that by the first decade of the new millennium "shock" appears in more than 23 percent of all articles in economics. Year-by-year analysis of the 1960s and 1970s localizes the take-off point to 1973. The gray line, which presents the share of all articles that mention a family of terms identifying macroeconomic articles and also "shock" or "shocks," is even more striking.[2] It lies somewhat above the black line until the 1960s. It takes off at the same point but at a faster rate, so that, by the first decade of the millennium, "shock" appears in more than 44 percent of macroeconomics articles.

Figure 1
Usage for "Shock"



[2] The macroeconomics family includes "macroeconomic," "macroeconomics," "macro economic," "macro economics," and "monetary." Because the search tool in *JSTOR* ignores hyphens, this catches the common variant spellings, including hyphenated spellings.

Since the 1970s the macroeconomic debate has been centered to some extent on shocks: the divide between real business cycle theorists and new Keynesian macroeconomists evolved around the importance of real versus nominal shocks for business cycle fluctuations. More important, shocks became a central element in observing the macroeconomic phenomena. Then, one question to be addressed in this paper is, how can we account for that terminological transformation? Our answer consists of a story about how the meaning of "shock" became sharpened and how shocks themselves became the objects of economic observation – both shocks as phenomena that are observed using economic theory to interpret data and shocks themselves data that become the basis for observing phenomena, which were not well articulated until shocks became observable. Here we are particularly interested in the debates carried out in the macroeconomic literature of the business cycle.

Among economists "shock" has long been used in a variety of ways. The earliest example in *JSTOR* underscores a notion of frequent but irregular blows to the economy: "an unending succession of slight shocks of earthquake to the terrestrial structure of business, varying of course in practical effect in different places and times …" (Horton 1886, 47). Near the same time, we also find the idea that shocks have some sort of propagation mechanism:

> Different industrial classes have in very different degrees the quality of economic elasticity; that is, the power of reacting upon and transmitting the various forms of economic shock and pressure. [Giddings (1887), 371]

Francis Walker (1887, 279) refers to "shocks to credit, disturbances of production, and fluctuations of prices" as a cause of suffering, especially among the working classes. Frank Taussig (1892, 83) worries about a "shock to confidence" as a causal factor in the price mechanism. Charles W. Mixter (1902, 411) considers the transmission of the "shock of invention" to wages. Among these early economists, the metaphor of shocks may refer to something small or large, frequent or infrequent, regularly transmissible or not. And while these varieties of usages continue to the present day, increasingly shocks are regarded as transient features of economic time series subject to well-defined probability distributions, transmissible through regular deterministic processes. Over time, shocks have come to be regarded as the objects of economic analysis and, we suggest, as *observable*.

What does it mean to be observable? The answer is often merely implicit – not only among economists, but among other scientists. Critics of scientific realism typically take middle-sized entities (for example, tables and chairs) as unproblematically observable. They object to the claims of scientific realists for the existence of very tiny entities (e.g., an electron) or very large entities (e.g., the dark matter of the universe) in part on the grounds that they are not observable, taking them instead to be theoretical constructions that may or may not really exist. Realist philosophers of science also accept everyday observation as unproblematic, but may respond to the critics by arguing that instruments such as microscopes, telescopes, and cloud chambers are extensions of our ordinary observational apparatus and that their targets are, in fact, directly observed (and are not artifacts of these apparatuses). The critics point out that substantial theoretical commitments are involved in "seeing" inaccessible entities with such instruments, and what we see would change if we rethought those commitments – hardly the mark of something *real* in the sense of independent of ourselves and our own thinking.

In a contribution to this debate that will form a useful foil in our historical account, the philosophers James Bogen and James Woodward argue that we ought to draw a distinction between *data* and *phenomena*:

> Data, which play the role of evidence for the existence of phenomena, for the most part can be straightforwardly observed. However, data typically cannot be predicted or systematically explained by theory. By contrast, well-developed scientific theories do predict and explain facts about phenomena. Phenomena are detected through the use of data, but in most cases are not observable in any interesting sense of that term. [Bogen and Woodward (1988), 305-306].

Cloud-chamber photographs are an example of data, which may provide evidence for the phenomena of weak neutral currents. Quantum mechanics predicts and explains weak neutral currents, but not cloud chamber photographs.

Qualifiers such as "typically," "in most cases," and "in any interesting sense" leave Bogen and Woodward with considerable wiggle room. But we are not engaged in a philosophical investigation *per se*. Their distinction between data and phenomena provides us with a useful framework for discussing the developing epistemic status of shocks in (macro)economics.

We can see immediately that economics may challenge the basic distinction at a fairly low level. The U.S. Bureau of Labor Statistics collects information on the prices of various goods in order to construct the consumer price index (CPI).[3] Although various decisions have to be made about how to collect the information – for example, what counts as the same good from survey to survey or how to construct the sampling, issues that may be informed by theoretical considerations – it is fair to consider the root information to be data in Bogen and Woodward's sense. These data are transformed into the price indices. The construction of index numbers for prices has been the target of considerable theoretical discussion – some purely statistical, some drawing on economics explicitly. Are the price indices, then, phenomena – not observed, but explained by theory? The theory used in their construction is not of an explanatory or predictive type; rather it is a theory of measurement – a theory of how to organize information in a manner that could be the target of an explanatory theory or that may be used for some other purpose. We might, then, wish to regard – as economists almost always do – the price indices as data. But can we say that such data are "straightforwardly observed"?

The raw information from which the price indices are constructed also fits Bogen and Woodward's notion that the object of theory is not to explain the data. For example, the quantity theory of money aims to explain the price level or the rate of inflation, but not the individual prices of gasoline or oatmeal that form part of the raw material from which the price level is fabricated. While that suggests that the price level is a phenomenon, here again we might question whether the object of explanation is truly the price level (say a specific value for the CPI at a specific time) or whether it is rather the fact of the proportionality of money and prices, so that the observations of the price level are data, and the proportionality of prices to money is the phenomenon.

The ambiguity between data and phenomena, an ambiguity between the observable and the inferred, is recapitulated in the ambiguity in the status of shocks, which shift from data to phenomena and back – from observed to inferred – depending on the target of theoretical explanation. Our major goal is to explain how the changing epistemic status of shocks and the changing understanding of their observability accounts for the massive increase in their role in economics as documented in Figure 1. Shocks moved from secondary factors to the center of economic attention after 1973.

[3] See Boumans (2005, ch. 6) and Stapleford (2009) for discussions of the history and construction of index numbers.

That story tells us something about economic observation. And even though we are telling an historical tale, and not a methodological or philosophical one, the history does, we believe, call any simple reading of Bogen and Woodward's distinction between data and phenomena into question and may, therefore, be of some interest to philosophers as well as historians.

II. The New Classical Macroeconomics and the Rediscovery of Shocks

Although shock continues to be used with a wide range of meanings, after 1973 the idea of shocks as pure transients or random impulses conforming to a probability distribution or the same random impulses conforming to a time-series model independent of any further economic explanation became dominant. Why? Our thesis is that it was the inexorable result of the rise of the new classical macroeconomics and one of its key features the rational expectations hypothesis, originally due to John Muth (1961) but most notably promoted in the early macroeconomic work of Robert Lucas (e.g., Lucas 1972) and Thomas Sargent (1972).

While rational expectations has been given various glosses (for example, people use all the information available or people know the true model of the economy), the most relevant one is probably Muth's original statement: "[rational] expectations . . . are essentially the same as the predictions of the relevant economic theory" (Muth 1961, 315, 316). Rational expectations on this view are essentially equilibrium or consistent expectations. A standard formulation of rational price expectations (e.g., in Hoover 1988, 187) is that $p_t^e = E(p_t \mid \Omega_{t-1})$, where p is the price level, Ω is *all* the information available in the model, t indicates the time period, e indicates an expectation, and E is the mathematical conditional expectations operator. The expected p_t^e can differ from the actual price p_t but only by a mean-zero, independent, serially uncorrelated random error. The feature that makes the expectation an equilibrium value analogous to a market clearing price is that the content of the information set Ω_{t-1} includes the model itself, so that an expected price would not be consistent with the information if it differed from the best conditional forecast using the structure of the model, as well as the values of any exogenous variables known at time $t - 1$.

The mathematical expectations operator reminds us that "to discuss rational expectations formation at all, some explicit stochastic description is clearly required" (Lucas 1973, 328-329, fn. 5). Yet, the need for a regular, stochastic characterization of the impulses to the economy places a premium on shocks with straightforward time-series representations; and this meaning of shock increasingly became the dominant one. The same pressure that led to the characterization of shocks as the products of regular, stochastic processes also suggested that government policy be characterized similarly – that is, by a policy rule with possibly random deviations. The economic, behavioral rationale was, first, that policymakers, like other agents in the economy, do not take arbitrary actions, but systematically pursue goals, and, second, that other agents in the economy anticipate the actions of policymakers.

Sargent relates the analysis of policy as rules under rational expectations to general equilibrium: "Since in general one agent's decision rule is another agent's constraint, a logical force is established toward the analysis of dynamic general equilibrium systems" (1982, 383). Of course, this is a model-relative notion of general equilibrium (that is, it is general only to the degree that the range of the conditioning of the expectations operator, $E(\cdot|\cdot)$, is unrestricted relative to the information set, Ω_{t-1}). Lucas took matters a step further taking the new technology as an opportunity to integrate macroeconomics with a version of the more expansive Arrow-Debreu general-equilibrium model. He noticed the equivalence between the intertemporal version of that model with contingent claims and one with rational expectations. In the version with rational expectations, it was relatively straightforward to characterize the shocks in a manner that reflected imperfect information – in contrast, to the usual perfect-information framework of the Arrow-Debreu model – and generated more typically macroeconomic outcomes. Shocks were a centerpiece of his strategy:

> viewing a commodity as a function of stochastically determined shocks . . . in situations in which information differs in various ways among traders . . . permits one to use economic theory to make precise what one means by information, and to determine how it is valued economically. [Lucas 1980, 707]

His shock-oriented approach to general-equilibrium models of business cycles was increasingly applied to different areas of macroeconomics.

Rational expectations, the focus on market-clearing, general-equilibrium models, and the characterization of government policy as the execution of stable rules came together in Lucas's (1976) famous policy noninvariance argument (the "Lucas critique"): if macroeconometric models characterize the time-series behavior of variables without explicitly accounting for the underlying decision problems of the individual agents who make up the economy, then when the situations in which those agents find themselves change, their optimal decisions will change, as will the time-series behavior of the aggregate variables. The general lesson was that a macroeconometric model fitted to aggregate data would not remain stable in the face of a shift in the policy rule and could not, therefore, be used to evaluate policy counterfactually.

In one sense, Lucas merely recapitulated and emphasized a worry that Haavelmo (1940) had already raised – namely, that a time-series characterization of macroeconomic behavior need not map onto a structural interpretation. But Haavelmo's (1944, ch. II, section 8) notion of structure was more relativized than the one that Lucas appeared to advocate. Lucas declared himself to be the enemy of "free parameters" and took the goal to be to articulate a complete general equilibrium model grounded in parameters governing "tastes and technology" and in exogenous stochastic shocks (1980, esp. 702 and 707). Lucas's concept of structure leads naturally to the notion that what macroeconometrics requires is microfoundations – a grounding of macroeconomic relationships in microeconomic decision problems of individual agents (see Hoover 2010). The argument for microfoundations was barely articulated before Lucas confronts its impracticality – analyzing the supposedly individual decision problems not in detail but through the instrument of "'representative' households and firms" (Lucas 1980, 711).

The Lucas critique stood at a crossroads in the history of empirical macroeconomics. Each macroeconometric methodology after the mid-1970s has been forced to confront the central issue that it raises. Within the new classical camp, there were essentially two initial responses to the Lucas critique – each in some measure recapitulating approaches from the 1930s through the 1950s.

Lars Hansen and Sargent's (1980) work on maximum-likelihood estimation of rational expectations models and subsequently Hansen's work on generalized method-of-moments estimators initiated (and exemplified) the first (conservative) response (Hansen 1982; Hansen and Singleton 1982). Hansen and Sargent attempted to maintain the basic framework of the Cowles-Commission program of econometric identification (inspired by Haavelmo 1944) in which theory provided the deterministic structure that allowed the error to be characterized by manageable probability distributions and thus set aside (Koopmans 1950; Hood and Koopmans 1953). The target of explanation remained – as it had been for Frisch, Tinbergen, Klein, and the large-scale macroeconometric modelers – the conditional paths of aggregate variables. The structure was assumed to be known *a priori* and measurement was directed to the estimation of parameters, now assumed to be "deep" – at least relative to the underlying representative-agent model.

Finn Kydland and Edward Prescott, starting with their seminal real-business cycle model in 1982, responded with a radical alternative to Hansen and Sargent. Instead of embracing, they rejected the soundness of Haavelmo's division of labor between economic and statistical theories (see Kydland and Prescott 1990, 1991, esp. 164-167; Prescott 1986; Hoover 1995, 28-32).

Though neither Haavelmo nor his followers in the Cowles Commission clearly articulated either the fundamental nature of the *a priori* economic theory that was invoked to do so much work in supporting econometric identification or the ultimate sources of its credibility, Haavelmo's decomposition became the centerpiece of econometrics (being an unassailable dogma in some quarters).

Kydland and Prescott took the message from the Lucas critique that a workable model must be grounded in microeconomic optimization (or in as near to it as the representative-agent model would allow). And they accepted Lucas's call for a macroeconomic theory based in general equilibrium with rational expectations. Though they held these theoretical presuppositions dogmatically – propositions which were stronger and more clearly articulated than any account of theory offered by Haavelmo or the Cowles Commission – they also held that models were at best workable approximations and not detailed, "realistic" recapitulations of the world. Thus, they rejected the Cowles Commission's notion that the economy could be so finely recapitulated in a model that the errors could conform to a tractable probability law and that its true parameters could be the objects of observation or direct measurement.

Having rejected Haavelmo's "probablity approach," their alternative approach embraced Lucas's conception of models as simulacra:

> . . . a theory is . . . an explicit set of instructions for building a parallel or analogue system – a mechanical, imitation economy. A good model, from this point of view, will not be exactly more real than a poor one, but will provide better imitations.
>
> . . .
>
> Our task . . . is to write a FORTRAN program that will accept specific economic policy rules as "input" and will generate as "output" statistics describing the operating characteristics of time series we care about, which are predicted to result from these policies. [Lucas 1980, 696-697, 709-710]

On Lucas's view, a model needed to be realistic only to the degree that it captured some set of key elements of the problem to be analyzed and successfully mimicked economic behavior on those limited dimensions. Given the preference for general-equilibrium models with few free parameters, shocks in Lucas's framework became the essential driver and the basis on which models could be assessed: "we need to test [models] as useful imitations of reality by subjecting them to shocks for which we are fairly certain how actual economies, or parts of economies, would react" (Lucas 1980, 697).

Kydland and Prescott, starting with their first real-business-cycle model (1982), adopted Lucas's framework. Real (technology) shocks were treated as the main driver of their model and its ability to mimic business-cycle phenomena when shocked became the principal criterion for the empirical success (Prescott 1986; Kydland and Prescott 1990, 1991; and Kehoe and Prescott 1995). Shocks in Lucas's and Kydland and Prescott's framework assumed a new and now central crucial task: they became the instrument through which the real-business-cycle modeler would select the appropriate artificial economy to assess policy prescriptions. For this, it is necessary to identify correctly substantive shocks – that is, the ones the effect of which on the actual economy could be mapped with some degree of confidence. Kydland and Prescott's translation of Lucas's conception of modeling into the real-business-cycle model generated a large literature.

Both Kydland and Prescott's earliest business cycle model as well their successors, the so-called *dynamic stochastic general-equilibrium* (*DSGE*) *models*, were developed explicitly within Lucas's conceptual framework, though they subsequently were adopted by economists with quite different methodological orientations. Kydland and Prescott (1982) presented a tightly specified, representative-agent, general-equilibrium model in which the parameters were calibrated. They rejected statistical estimation – specifically rejecting the Cowles Commission's "systems-of-equations" approach – on the ground that Lucas's conception of modeling required matching reality only on specific dimensions and that statistical estimation penalized models for not matching it on dimensions that in fact were unrelated to "the operating characteristics of time series we care about." *Calibration* involves drawing parameter values from general economic considerations: both long-run unconditional moments of the data and facts about national-income accounting, as well as evidence from independent sources, such as microeconomic studies (Kydland and Prescott 1996, 74).

To evaluate their model, Kydland and Prescott (1982) adopt the "test of the Adelman's," which is essentially a Turing test: would a business-cycle analyst be unable to distinguish the artificial output of a model from the data on the actual economy (Adelman and Adelman 1959; King and Plosser 1989; Kydland and Prescott 1990, 6; see also Lucas 1977, 219, 234)? Kydland and Prescott's main criterion is how well the unconditional second moments of the simulated data matched the same moments in the real-world data. To generate the simulation, they simply drew shocks from a probability distribution the parameters of which were chosen to ensure that the variance of output produced in the model matched exactly the corresponding value for the actual U.S. economy (Kydland and Prescott 1982, 1362). This, of course, was a violation of Lucas's maxim: do not rely on free parameters. Given that shocks were not, like other variables, supplied in government statistics, their solution in later work as to take the "Solow residual" as the measure of technology shocks. In effect, they used the production function as an instrument to measure technology shocks (Prescott 1986, 14-16)

Kydland and Prescott treated the technology shocks measured by the Solow residual as data in Bogen and Woodward's sense. As with price indices, certain theoretical commitments were involved. Prescott (1986, 16-17) discussed various ways in which the Solow residual may fail accurately to measure true technology shocks, but concluded that, for the purpose at hand, that they would serve adequately. The key point at this stage is that – in keeping with Bogen and Woodward's distinction – Kydland and Prescott were not interested in the shocks *per se*, but in what might be termed "the technology-shock phenomenon." The Solow residual is serially correlated. Prescott (1986, 14 and 15, fn. 5) treated it as governed by a time-series process. He claimed that very similar simulations and measures of business-cycle phenomena (that is, of the cross-correlations of current GDP with a variety of variables at different lags) would result whether the shocks were modeled as nonstationary or as stationary but highly persistent (see also Kydland and Prescott 1990).

Kydland and Prescott's simulations were not based on direct observation of technology shocks (that is, on the Solow residual), but on the statistical characterization of those shocks (the technology-shock phenomenon). The earlier simulation studies of Adelman and Adelman (1959) had been concerned not with the shocks, but with the time paths of variables: the shock phenomenon was thus *secondary*. But for Kydland and Prescott, who focused on the covariation of the variables rather than their time paths, technology-shock phenomenon was *primary*.

In contrast to the Adelmans, whose measures of shocks depended on the whole structure of the model, Kydland and Prescott's technology shocks were measured by just one element of the model, the Cobb-Douglas production function. Measured this way, technology shocks on Kydland and Prescott's view have a degree of model-independence and an integrity that allows them to be transferred between modeling contexts.

Although real-business-cycle modelers typically use technology shocks to characterize the shock process, the technology-shock phenomenon, they have from time to time treated them as direct inputs into their models (essentially as observed data). Hansen and Prescott (1993) fed technology shocks directly into a real-business-cycle model in order to model the time path of U.S. GDP over the 1990-91 recession.[4]

III. The Identification of Shocks

Whereas Kydland and Prescott had attacked Haavelmo's and the Cowles Commission's assumption that models define a tractable probability distribution, Sims attacked the credibility of the *a priori* assumptions that they used to identify the models (Sims 1980, 1, 2, 14, 33). Nonetheless, the positive contribution of Sims's approach that bears most strongly on our story. Sims asks – to quote the title of Sargent and Sims's (1977) earlier paper – what can be learned about business cycles "without pretending to have too much *a priori* economic theory"?

Sims's (1980) took general-equilibrium, in one sense, more seriously than did the Cowles Commission in that he treated all the independently measured economic variables as endogenous. Although, as with Haavelmo, Sims divided the model into a deterministic and an indeterministic part, he rejected the notion that the deterministic part was structural. He regarded his system of equations – the *vector-autoregression* (*VAR*) *model* – as a reduced form in which the random residuals were now the *only* drivers of the dynamics of the model and, hence, considerably more important than they had been in the Cowles Commission's approach. Sims referred to these residuals as "innovations," which stressed the fact that they were independent random shocks without their own time-series dynamics. Since the deterministic part of the model was not structural, all time-series behavior could be impounded there, so the shocks are now pure transients.

Sims used his VAR model to characterize dynamic phenomena through variance decomposition analysis and impulse-response functions. Variance decomposition is an accounting exercise that determines the proportion of the

[4] The major reason for the focus of the real-business-cycle (RBC) literature on comparing unconditional moments is the way they characterized cycles as recurrent fluctuations in economic activity, going back to Mitchell and Burns (1946) through Lucas's equilibrium approach (Cooley and Prescott 1995, 26; Kydland and Prescott 1982, 1359-1360). King, Plosser and Rebelo (1988) provide an early example of a calibrated RBC model looking at time paths, while Christiano (1988) develops an estimated RBC model that compares theoretical and observed time paths.

variability of each variable that is ultimately attributable to the exogenous shocks to each variable. The impulse-response function traces the effect on the time-series for a variable from a known shock to itself or to another variable. Particular shocks need not be measured or observed in order to conduct either of these exercises; nonetheless, they must be characterized. The dynamics of the data must be cleanly divided between the deterministic part and the independent random shocks. The difficulty, however, is that, in general, there is no reason that the residuals to an estimated VAR ought to have the characteristic of independent random shocks – in particular, they will generally be correlated among themselves.

To deal with the problem of intercorrelated residuals, Sims assumed that the variables in his VAR could be ordered recursively (in a *Wold causal chain* via Cholesky decompositions, for example), in which a shock to a given variable immediately affects that same variable and all those lower in the system. The coefficients on the contemporaneous variables are selected so that the shocks are orthogonal to each other.

Sims (1980, 2) admitted "that the individual equations of the model are not products of distinct exercises in economic theory" – that is, not structural in the Cowles Commission's sense. And he suggested that "[n]obody is disturbed by this situation of multiple possible normalizations." In fact, given N variables, there are $N!$ possible normalizations (e.g., for $N = 6$, there are 720 normalizations). And far from nobody being disturbed, critics immediately pointed out that, first, the variance decompositions and the impulse-response functions were, in general, not robust to the choice of normalization; and, second, policy-analysis required not just one of the possible renormalizations, but the right one. Sims (1982, 1986) rapidly conceded the point. The VAR approach did not eliminate the need for identifying assumptions. Yet, Sims had nevertheless changed the game.

The Cowles Commission had sought to measure the values of structural parameters through imposing identifying assumptions strong enough to recover them all. Sims had shown that, if the focus of attention was on identifying the shocks themselves, then the necessary identifying assumptions were weaker: with a *structural VAR* (SVAR) – that is, a VAR with orthogonalized shocks – one needs to know only the recursive order or, more generally, the causal structure of the *contemporaneous* variables. The parameters of the lagged variables in the dynamic system need not be structural, so that the SVAR is a quasi-reduced form and less is taken on faith than in the Cowles Commission's or calibrationist frameworks.

The SVAR put shocks front and center – not because shocks could not have been identified in the Cowles Commission's framework nor because shocks are automatically interesting in themselves, but because the time-series properties of the shocks are essential to the identification strategy. Variance-decomposition exercises and impulse-response functions do not necessarily consider measured shocks, but rather ask a simple counterfactual question, "what would be the effect of a generic shock u of size v to variable x on variables x, y, and z?" The situation is essentially no different than that of technology shocks measured using the Solow residual. The SVAR, like a production function used to measure technology shocks, can be used as a measuring instrument to observed shocks to each variable in the VAR system. Just as the real-business-cycle modeler may be more interested in the generic business-cycle phenomena, so the SVAR modeler may be more interested in generic dynamic phenomena. But equally the SVAR modeler use the particular observed shocks to the whole system of equations to generate specific historical time paths for variables or to conduct counterfactual experiments (e.g., Sims 1999).

There are, however, key differences with the calibrationist approach. Calibrationists make very strong identifying assumptions with respect to structure.

Essentially, they claim to know not only the mathematical form of the economic relationships but their parameterization as well. The cost is that they give up on the notion that residuals will conform to tractable probability distributions. In contrast, the SVAR modeler makes minimal structural assumptions and specifies nothing about the values of the parameters other than that they must deliver orthogonal shocks. Whereas typical real-business-cycles are driven by technology shocks only, SVAR models necessarily observe shocks for each variable in the system.

IV. Coming Full Circle: Estimation by Impulse-Response Matching

Although starting from very different critical stances, both Sims's SVAR approach and the new classicals calibrationist approach elevated shocks to a starring role. Shocks had become the targets of measurement; models or parts of models had become the measuring instruments. In short, shocks were observable data in Bogen and Woodward's sense. Still, economists were frequently more interested in the phenomena that shocks generated – how the economy reacted generically to a particular type of shock – rather than in the particular shock to a particular variable on a particular date. Yet, the observability of shocks was *sine qua non* of identifying these phenomena in the first place. Both approaches, however, provided instances in which the particular values of shocks were treated as important in their own right.

Whether because of the similarity in their views of shocks or, perhaps, for the more mundane sociological reason that economists, like other scientists, cannot resist trying to make sense of each other's work and often seek out common ground, the 1990s witnessed a rapprochement between the DSGE and SVAR programs. Any DSGE model has a reduced-form representation, which can be seen as a special case of a more general VAR, and it also has a contemporaneous causal ordering of its variables that provides a basis for converting the VAR into an SVAR. A calibrated or estimated DSGE model, therefore, can generate variance decompositions and impulse-response functions, which may, in their turns, be compared directly to their counterparts generated from estimated SVARs in which DSGE models are nested. Such comparisons are methodologically equivalent to Kydland and Prescott's strategy of attempting to match the second moments of calibrated models to the equivalent statistics for actual data; they just use different target phenomena.

By the early 1990s the terms of the debate in macroeconomics had shifted from one between monetarists, such as Milton Friedman, and old Keynesians in the macroeconometric tradition, such as James Tobin and Lawrence Klein, or one between the old Keynesians and the new classicals into one between the *new Keynesians* and the new classicals (Hoover 1988, 1992). The new Keynesians essentially adopted the technical paradigms of the new classicals, typically including the rational-expectations hypothesis, but rejected the notion of perfect competition with continuous market clearing as a sound basis for macroeconomic models, which opened the door for activist policies to improve welfare. Sims (1989, 1992) regarded the debate between the new classicals – especially, the real-business-cycle modelers – and the new Keynesians as having reached an impasse. In his view, real-business-cycle modelers assessed their models with an impoverished information set (unconditional moments). Sims (1992, 980) argued that the debate between the monetarists and the old Keynesians had reached a similar impasse, which a focus on time-series information (mainly responses to innovations and Granger causality) had helped resolve by establishing that monetary policy has substantial effects on real output. Analogously, Sims (1992, 980) suggested that real-business-cycle modelers should consider the richer set of time-series information. He set out to provide real-business-cycle modelers to confront their models with "the documented impulse response facts about interactions of monetary and real variables" (980).

Sims wanted to reestablish the relevance of estimation methods in an area of research that had become dominated by calibration techniques, and he sought common ground in what amounted to adopting Lucas's views on modeling: to select a substantive shock and compare models by the implied dynamic responses to it; a good model is one in which the impulse-response function of the model matches the impulse-response function of the data, as determined through the instrumentality of the SVAR (see also Christiano 1988 and Singleton 1988). Once again, shocks were data used to characterize phenomena, and models were judged by their ability to reproduce those phenomena.

Sims's proposal must be distinguished from merely matching historical performance in the manner of Hansen and Prescott (1993). The interactions of the different elements are too complex to connect, for example, policy actions to particular outcomes (Leeper, Sims and Zha 1996, 2). Christiano, Eichenbaum and Evans (1999, 68), for example, argued the covements among aggregate variables cannot be interpreted as evidence for or against the neutrality of money, since a "given policy action and the economic events that follow it reflect the effects of *all* the shocks to the economy." Sims's proposal, following Lucas, amounted to a highly restricted counterfactual experiment in which the effects of an isolated shock can be traced out in the economy (that is, in the SVAR) and compared the analogous effects in a model. The goal was precisely analogous to experimental controls in a laboratory in which the effect of a single modification is sought against a stable background.

Much of the research in this vein focused on monetary shocks – that is, to shocks to short-term interest rates. The short-term interest rate was regarded as the central bank's policy instrument and assumed in the theoretical models to be governed by a policy rule – the central bank's *reaction function* (usually a "Taylor rule"). Monetary policy was, of course, an intrinsically interesting and important area of research. It also held out the promise of clearer discrimination among theoretical models "because different models respond very differently to monetary policy shocks" (Christiano, Eichenbaum and Evans 1999, 67).

A case that illustrates very clearly Sims's strategy is the so-called "price puzzle" (see Eichenbaum's 1992 comments on Sims 1992). Simple textbook models suggest that tighter monetary policy should reduce the rate of inflation and the price level. One might expect, therefore, that an exogenous positive shock to the short-term interest rate would result in an declining impulse-response-function for prices. In fact, Sims and most subsequent researchers found that the impulse-response function for prices in an SVAR tends to rise for some time before falling. The quest for a theoretical model that accounts for this robust pattern has generated a large literature (see Demiralp *et al.* 2010).

Sims's (1992) call for macroeconomists seriously to consider time-series evidence was taken into consideration subsequently. Whereas in his 1992 article he reported several point-estimate impulse response functions obtained from alternative VARs for data from different countries, Leeper, Sims and Zha (1996) focused on the U.S. data and used sophisticated VAR methods to characterize features of aggregate time-series data. Here, in contrast to Sims (1992), the authors present confidence intervals for the estimated impulse response functions (cf. Christiano, Eichenbaum and Evans 1996, 1999).

Parallel to characterizing dynamic responses to shocks in the data through VARs there was the effort of building artificial economies, small-scale dynamic general-equilibrium monetary models, to explain the business cycle phenomena and to derive policy implications of them. Sims himself joined this enterprise with Eric

Leeper (Leeper and Sims 1994; see also Christiano and Eichenbaum (1995), Yun (1996), and Christiano, Eichenbaum and Evans (1997)). Here the parameters were either estimated with methods such as maximum likelihood or general methods of moments, or were calibrated. Once the parameters were assigned numerical values, one can derive the theoretical impulse response functions to a monetary shock. However, the closeness of the match between the model-based and the SVAR-based impulse-response functions is ususally judged in a rough-and-ready fashion – the same ocular standard applied in matching unconditional moments in the real-business-cycle literature.

Rotemberg and Woodford (1997) and the literature that derived from this work took impulse-response matching one step further. Setting aside some of the fine details, the essence of their approach was to select the parameterization of the theoretical model in order to mimimize the distance between the impulse-response functions of the model and those of the SVAR, which became a standard approach in DSGE macroeconomics (only parameters that were identifiable were estimated, the others were calibrated). But Rotemberg and Woodford's (1997) model failed to deliver the slow responses ("intertia") observed in impulse-response functions generated from SVARs. Other economists took on the task of building DSGE models, estimated by impulse-response matching that captured the inertia of the impulse-response functions (Christiano, Eichenbaum and Evans 2005 and Smets and Wouters 2007 – see Duarte forthcoming).

Rottemberg and Woodford's method, in effect, treated the impulse-response functions of the SVAR as data in their own right – data that could be used as an input to the estimator. Where previously, the shock could be regarded as data and the impulse-response functions phenomena, the shocks were now moved down a level. They stood in the same relationship to the new data as the raw prices of individual goods did the price index. And the focus of the technique shifted from the isolation of shocks and mimicking of dynamic phenomena back, as it had in the post-Cowles Commission macroeconometric program, to the measurement of structural parameters.

V. Shocks, Macroeconometrics, and Observability

We have addressed three main questions in this paper. Two were explicit: What is the relationship of shocks to observation? Why did the uses of the language of shocks explode after the early 1970s? And one question was only implicit: What lessons does the history of shocks provide to philosophers of science or economic methodologists? The answers to these three questions are deeply entangled in our narrative.

In the earliest days of modern econometrics in the 1930s, estimated equations were conceived of as having unobservable error terms. Yet, these systems of equations, which had their own deterministic dynamics were also thought of as being perturbed by actual disturbances, so that the error terms were – to use Frisch's terminology – a mixture of stimuli and aberration. Business-cycle theory was principally interested in the stimuli. Business-cycle theory gave way after World War II to a theory of macroeconomic policy that aimed to avoid cycles in the first place. Attention thus shifted to the deterministic parts of structural models and, notwithstanding Haavelmo's characterization of shocks as well-behaved phenomena with a regular probabilistic structure, shocks became of secondary interest.

It was only when the introduction of the rational expectations hypothesis compelled economists to treat the stochastic specification of a model as a fundamental element rather than a largely ignorable supplement that shocks returned to center stage and economists began to notice that models could be treated as measuring instruments through which shocks became observable. Rational expectations compel at least a relative-to-modeled-information general-equilibrium approach to modeling. Thoroughly done, such an approach – whether theoretically, as in a real-business-cycle model, or econometrically, as in an SVAR – endogenizes every variable except the shocks. Shocks are then elevated to be the sole drivers of economic dynamics, and their observability, if not their particular values, becomes the *sine qua non* of a properly specified model. It is, therefore, hardly surprising that a vast rise in the usage of shock occurs after 1973, since shocks are central to a fundamental reconceptualization of macroeconomic theory that, to be sure, began with Frisch forty years earlier, but did not sweep the boards until the rise of the new classical macroeconomics.

We have used Bogen and Woodward's distinction between observable data and inferred phenomena to provide an organizing framework for our discussion. Although it may prove useful as a rough-and-ready contrast, it appears not to draw a bedrock distinction: at some points shocks could be best regarded as phenomena, inferred from observable data; at other points as data observed using models as measuring instruments; or as the raw material from which data were constructed and which were then used as an input to generate further phenomena or as the basis for higher-order inference. Economics, even in its deepest reaches, is about relationships. What the history of shocks shows is that when we give up the rather tenuous grounding of observability in human senses, then the distinctions between observable and inferrable and between data and phenomena are, at best, relative ones that depend on our principal interests and our targets of explanation, on our presuppositions, explicitly theoretical or merely implicit, and on the modeling tools we have at our disposal – which emphasizes the role of models as measuring instrument (Boumans 2005, esp. 16-17) that integrate a range of ingredients coming from disparate sources, and as autonomous agents that mediate theories and the real world (Morgan and Morrison 1999). Philosophers of science would do well to consider such cases.

References

Adelman, Irma and Frank L. Adelman. (1959) "The Dynamic Properties of the Klein-Goldberger Model," *Econometrica* 27(4):596-625.

Bogen, James and James Woodward (1988) "Saving the Phenomena," *Philosophical Review* 97(3):303-352.

Boumans, Marcel. (2005). *How Economists Model the World into Numbers*. London: Routledge.

Christiano, Lawrence J. (1988). "Why Does Inventory Investment Fluctuates So Much?" *Journal of Monetary Economics*, 21(2/3):247-280.

Christiano, Lawrence J. and Martin Eichenbaum (1995). "Liquidity Effects, Monetary Policy, and the Business Cycle," *Journal of Money, Credit and Banking*, 27(4):1113-1136

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans (1996). "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," *Review of Economics and Statistics*, 78(1):16-34

________, ________, and ________ (1997). “Sticky price and limited participation models of money: a comparison,” *European Economic Review,* 41(6):1201-1249.

________, ________, and ________ (1999) “Monetary Policy Shocks: What Have We Learned and to What End?,” in John Taylor and Michael Woodford (eds.) *Handbook of Macroeconomics* Amsterdam: North-Holland, pp. 65-148.

________, ________, and ________ (2005). “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,” *Journal of Political Economy*, 113(1):1-45.

Cooley, Thomas F. and Edward C. Prescott (1995). “Economic Growth and Business Cycles,” in Cooley and Prescott (eds.). *Frontiers of Business Cycle Research*. Princeton: Princeton University Press.

Demiralp, Selva, Kevin D. Hoover, and Stephen J. Perez. (2010) “Still Puzzling: Evaluating the Price Puzzle in an Empirically Identified Structural Vector Autoregression,” unpublished working paper, Duke University.

Duarte, Pedro Garcia (forthcoming). “Recent Developments in Macroeconomics: The DSGE Approach to Business Cycles in Perspective,” in Wade Hands and John Davis (Eds.), *The Elgar Companion to Recent Economic Methodology*. Edward Elgar.

Eichenbaum, Martin. (1992) “Comments ‘Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy’,” *European Economic Review* 36(5):1001-1011.

Frisch, Ragnar. (1933) “Propagation Problems and Impulse Problems in Dynamic Economics,” in *Economic Essays in Honor of Gustav Cassel*. London: George Allen and Unwin, pp. 171-205.

Giddings, Franklin H. (1887) “The Theory of Profit-Sharing,” *Quarterly Journal of Economics* 1(3):367-376.

Haavelmo, Trygve. (1940) “The Inadequacy of Testing Dynamic Theory by Comparing Theoretical Solutions and Observed Cycles,” *Econometrica* 8(4):312-321.

________ (1944) “The Probability Approach in Econometrics,” *Econometrica* 12 (supplement), July.

Hansen, Gary D. and Edward C. Prescott. (1993) “Did Technology Shocks Cause the 1990-1991 Recession?” *American Economic Review* 83(2):280-286.

Hansen, Lars Peter. (1982) “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica* 50(4):1029-1054.

Hansen, Lars Peter and Thomas J. Sargent. (1980) “Formulating and Estimating Dynamic Linear Rational Expectations Models,” in Robert E. Lucas, Jr. and Thomas J. Sargent (eds.) *Rational Expectations and Econometric Practice*. London: George Allen & Unwin, 1981.

Hansen, Lars Peter and Kenneth J. Singleton. (1982) "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica* 50(5):1269-1286.

Hood, W.C. and T.C. Koopmans, editors. (1953) *Studies in Econometric Method*. New York: Wiley.

Hoover, Kevin D. (1988) *The New Classical Macroeconomics: A Sceptical Inquiry*. Oxford: Blackwell.

________ (1992) "The Rational Expectations Revolution," *Cato Journal* 12(1):81-106.

________. (1995) "Facts and Artifacts: Calibration and the Empirical Assessment of Real-Business-Cycle Models," *Oxford Economic Papers* 47(1):24-44.

________. (2010) "Microfoundational Programs," unpublished working paper, Duke University.

Horton, S. Dana. (1886) "Silver Before Congress in 1886," *Quarterly Journal of Economics* 1(1):45-75.

Kehoe, Timothy J., and Edward C. Prescott (1995). "Introduction to the Symposium: The Discipline of Applied General Equilibrium." *Economic Theory*, *6*(1): 1-11.

King, Robert G. and Charles I. Plosser. (1989) "Real Business Cycles and the Test of the Adelmans," *Journal of Monetary Economics* 33(2):405-438.

King, Robert G., Charles I. Plosser and Sergio Rebelo (1988). "Production, Growth and Business Cycles: I. The Basic Neoclassical Model," *Journal of Monetary Economics*, 27(2/3):195-232.

Koopmans, Tjalling C. (1950) *Statistical Inference in Dynamic Economic Models*. New York: Wiley.

Kydland, Finn E. and Edward C. Prescott. (1982) "Time to Build and Aggregate Fluctuations," *Econometrica* 50(6): 1345-1370.

________, and ________ (1990) "Business Cycles: Real Facts and a Monetary Myth," *Federal Reserve Bank of Minneapolis Quarterly Review* 14(2):3-18.

________, and ________ (1991) "The Econometrics of the General Equilibrium Approach to Business Cycles," *Scandinavian Journal of Economics*, 93(2):161-78.

________, and ________ (1996) "The Computational Experiment: An Econometric Tool," *Journal of Economic Perspectives*, 10(1):69-85.

Leeper, Eric, Christopher A. Sims (1994). "Toward a Modern Macroeconomic Model Usable for Policy Analysis," *NBER Macroeconomics Annual*, 9:81-118.

Leeper, Eric M., Christopher A. Sims, and Tao Zha. (1996) "What Does Monetary Policy Do? *Brookings Papers on Economic Activity*, no. 2:1-78.

Lucas, Robert E., Jr. (1972) "Econometric Testing of the Natural Rate Hypothesis," reprinted in Lucas (1981), pp. 90-103.

________ (1973) "Some International Evidence on Output-Inflation Tradeoffs," *American Economic Review*, 63(3):326-334.

________. (1976) "Econometric Policy Evaluation: A Critique," reprinted in Lucas (1981), pp. 104-130.

________. (1977) "Understanding Business Cycles," reprinted in Lucas (1981), pp. 215-240.

________. (1980). "Methods and Problems in Business Cycle Theory," *Journal of Money, Credit and Banking*, 12(4, Part 2):696-715.

________. (1981) *Studies in Business Cycle Theory*. Oxford: Blackwell.

Mitchell, Wesley Clair and Arthur F. Burns. (1946) *Measuring Business Cycles*. New York: National Bureau of Economic Research.

Mixter, Charles W. (1902) "Böhm-Bawerk on Rae," *Quarterly Journal of Economics*, 16(3):385-412.

Morgan, Mary and Margaret Morrison (1999). "Models as Mediating Instruments," in Morgan and Morrison (eds.), *Models as Mediators*. Cambridge: Cambridge University Press.

Muth, John F. (1961) "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29(3):315-335.

Prescott, Edward C. (1986) "Theory Ahead of Business Cycle Measurement," *Federal Reserve Bank of Minneapolis Quarterly Review*, 10(4):9-22.

Rotemberg, Julio and Michael Woodford (1997). "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," *NBER Macroeconomics Annual*, 12:297-346.

Sargent, Thomas J. (1972) "Rational Expectations and the Term Structure of Interest Rates," *Journal of Money, Credit and Banking*, 4(1, Part 1):74-97.

Sargent, Thomas J. and Christopher A. Sims. (1977) "Business Cycle Modeling Without Pretending to Have Too Much *A-priori* Economic Theory," in Sims *et al.*, editors. *New Methods in Business Cycle Research*. Minneapolis: Federal Reserve Bank of Minneapolis.

Sims, Christopher A. (1980) "Macroeconomics and Reality," *Econometrica*, 48(1):1-48.

________ (1982) "Policy Analysis with Econometric Models," *Brookings Papers on Economic Activity*, 13(1):107-52.

________ (1986) "Are Forecasting Models Usable for Policy Analysis?" *Federal Reserve Bank of Minneapolis Quarterly Review*, 10:2-15.

________ (1989) "Models and Their Uses," *American Journal of Agricultural Economics*, 71(2):489-94.

________ (1992). "Interpreting the Macroeconomic Time Series Facts – The Effects of Monetary Policy," *European Economic Review*, 36(5):975-1000.

________ (1999) "The Role of Interest Rate Policy in the Generation and Propagation of Business Cycles: What Has Changed Since the '30s?" in Jeffrey C. Fuhrer and Scott Schuh, editors., *Beyond Shocks: What Causes Business Cycles*. Federal Reserve Bank of Boston Conference Series, No. 42. Boston: Federal Reserve Bank of Boston, pp. 121-60.

Singleton, Kenneth J. (1988). "Econometric Issues in the Analysis of Equilibrium Business Cycle Models," *Journal of Monetary Economics*, 21(2/3):361-386.

Smets, Frank and Raf Wouters (2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3):586-606.

Stapleford, Thomas. (2009) *The Cost of Living in America: A Political History of Economic Statistics, 1880–2000*. South Bend, IN: University of Notre Dame Press.

Taussig, Frank. W. (1892) "The Silver Situation in the United States," *Publications of the American Economic Association*, 7(1):7-118.

Walker, Francis A. (1887) "The Source of Business Profits," *Quarterly Journal of Economics*, 1(3):265-288.

Yun, Tack (1996). "Nominal price rigidity, money supply endogeneity, and business cycles," *Journal of Monetary Economics*, 37(2):345-370.