ipea Institute for Applied Economic Research

Chapter title	CHAPTER 22 – ARE EVIDENCE-BASED MACROECONOMIC POLICIES POSSIBLE? THE DIFFICULT RELATIONSHIP BETWEEN MACROECONOMICS AND EMPIRICAL EVIDENCE
Author	Ronaldo Fiani
DOI	http://dx.doi.org/10.38116/978-65-5635-070-7/chapter22

Book title	PUBLIC POLICY AND USE OF EVIDENCE IN BRAZIL: CONCEPTS, METHODS, CONTEXTS AND PRACTICES
Editors	Natália Massaco Koga Pedro Lucas de Moura Palotti Janine Mello Maurício Mota Saboya Pinheiro
Volume	-
Series	-
City	Brasilia
Publisher	Institute for Applied Economic Research (Ipea)
Year	2024
Edition	-
ISBN	978-65-5635-070-7
DOI	http://dx.doi.org/10.38116/978-65-5635-070-7

CHAPTER 22

ARE EVIDENCE-BASED MACROECONOMIC POLICIES POSSIBLE? THE DIFFICULT RELATIONSHIP BETWEEN MACROECONOMICS AND EMPIRICAL EVIDENCE

Ronaldo Fiani[1]

In the last three decades, the methods and conclusions of macroeconomics have deteriorated to the point that much of the work in this area no longer qualifies as scientific research.[2]
Paul Romer

1 INTRODUCTION

The established framework for empirically evaluating macroeconomic policies involves the examination of theoretically determined relationships from econometric models. This framework was laid in the 1940s, with the work of the Cowles Commission for Economic Research (1950) in the United States, and became consolidated after World War II, with its diffusion throughout the United States and Western Europe.

Lucas (1976) critique would generate a shift in this framework, leading to the construction of macroeconometric models with deep parameters; that is, parameters that reflect the behavior of rational maximizing agents, not only in face of the possibilities of choice but also regarding the policies adopted, in the approach that became known as rational expectations.

Such an inflection would eventually result in the current trend of model development harshly criticized by Summers (1991) and Romer (2016). These two authors do not have the same theoretical ambition and, therefore, the same scope in terms of the academic repercussion of Lucas (1976) but focus precisely on evaluating the *practice* of production and analysis of empirical evidence through econometric models in macroeconomics.

1. Associate professor at the Department of Economics of the Federal University of Rio de Janeiro (UFRJ). E-mail: fiani@ie.ufrj.br.
2. Romer (2016, p. 1).

Thus, this chapter is organized into the following sections. After this introductory section, the second section discusses the evolution and role of econometric models in the first decades of the twentieth century as a tool for analyzing and comparing evidence-based macroeconomic policies. The third section discusses Lucas (1976) critique, the major theoretical challenge to this type of empirical approach. The fourth section discusses Summers' (1991) critique, which is mainly focused on the structural models that were developed after Lucas' (1976) critique. The fifth section discusses Romer's critique of more recent econometric models, whose complexity often hinders an accurate assessment of the value of their results, which is often worsened by the manipulation of parameters by the analyst. The conclusion section examines the possibilities for more evidence-based macroeconomics considering what has been discussed in the previous sections. It will be argued that it is possible to establish the outlines of *more evidence-based* macroeconomics, paraphrasing Julian Reiss (2008).

It is important to stress that we do not intend to undertake an exhaustive review of the use of evidence in macroeconomics, which would be impossible within the limits of this paper. We only intend to present a quick overview of the treatment of evidence in the field of macroeconomics, as an introduction to the debate on evidence-based macroeconomic policies in the country.

2 THE WIDESPREAD USE OF ECONOMETRIC MODELS IN THE EARLY POST-WAR PERIOD AS A TOOL FOR EVIDENCE-BASED MACROECONOMIC POLICIES

The origin of econometrics as a source of empirical evidence for macroeconomic policymaking and evaluation can be traced to the work of the Cowles Commission for Economic Research. This commission was created in 1932, when Alfred Cowles, president of Cowles and Company, an investment consulting firm in Colorado Springs, began a survey of the accuracy of stock market experts over 1928-1932. This survey sparked Alfred Cowles' interest in basic economic research, which led him to offer financial support for the creation of the commission and to always shoulder most of the funding (Christ, 1952, p. 3).

Mathematician Charles F. Roos was the first research director of the Cowles Commission, and his book published in 1934, *Dynamic economics: theoretical and statistical studies of demand, production and prices*, the first in the series of Cowles Commission monographs that would play a major role in the development of econometrics for the next decades, included topics such as the demand for consumer goods, automobile demand for gasoline, demand for agricultural products, demand for capital goods, among others. Roos' book, however, neglected the problem of identification, so it was not possible to tell whether the author had estimated a demand curve or some linear combination of demand and sup-

ply functions (Dimand, 2019, p. 3). Only the second generation of the Cowles Commission, with names like Jacob Marschak and Tjalling Koopmans would be able to tackle the identification problem.

Interestingly, the concerns of the first generation of the Cowles Commission, which included names such as Alfred Cowles himself, mathematician Harold Thayer Davis, and Charles F. Roos, were generally very different from what would later become the research standard in econometrics. In this first stage, there was a concern with predicting changes in stock prices – notably Alfred Cowles himself – and the study of business cycles.[3]

As far as cycle analysis is concerned, a prominent role has been played by Harold T. Davis, who, ironically – when considering later developments – dismissed Keynes' general theory in a surprisingly superficial way (in a footnote) but showed great interest in Stanley Jevons' theory about the influence of sunspots on business cycles (Dimand, 2019, p. 4).

Eugene Slutsky – who did not participate in the Cowles Commission – played a key role in taking the focus of interest of the commission away from the study of cycles, starting with the translation of his paper *The summation of random as the source of cyclic processes* – originally published in Moscow – in the journal *Econometrica*, due to his methodological criticism of the statistical methods used in the analysis of cycles (Dimand, 2019, p. 5-6).[4] It was a rare case of solving empirical controversies in macroeconomics, in which the methodological discussion shifted the focus of theoretical interest.

This focus on cycles would be definitively abandoned when Jacob Marschak took over as research director of the Cowles Commission in 1943. His actions would promote an important change in the commission's research lines, establishing econometric studies as the main method of empirical research in macroeconomics. With the appointment of Marschak as research director, one of the central concerns of the commission would be the study of the statistical properties of estimating simultaneous equations with random errors (Dimand, 2019, p. 8), influenced by the works of Trygve Haavelmo,[5] Leonid Hurwicz[6] and Tjalling Koopmans.[7]

3. Charles Roos and Harold Davis, research directors of the Cowles Commission at its inception, before the move from Colorado Springs to the University of Chicago in 1939, were mathematicians interested in curve fitting and techniques for decomposing time series into i) trends; ii) multiple coincident cycles with different periodicities and amplitudes; and iii) erratic movements (Dimand, 2019, p. 7).

4. Eugene Slutsky was the first teacher of Jacob Marschak – then going by Jakob – in Kyiv before the First World War. In his critique, Slutsky pointed to the fact that the techniques employed by the commission's cycle analysts generated apparent cycles, even though there was no cycle in the original data (Dimand, 2019, p. 5-6).

5. *The statistical implications of a system of simultaneous equations*, published in the journal *Econometrica* in 1943, and *The probability approach in econometrics*, also published in *Econometrica* in 1944 (Dimand, 2019, p. 8).

6. *Stochastic models of economic fluctuations*, published in *Econometrica* in 1944 (Dimand, 2019, p. 8).

7. The dissertation *Linear regression analysis of economic time series*, of 1936, by Tjalling Koopmans, was published in the following year (Dimand, 2019, p. 8).

The turning point would be the Cowles Commission conference in Chicago through January and early February 1945, which, according to Edmond Malinvaud (1983, p. 7), would become the most influential conference on statistical inference ever held. Malinvaud (1983, p. 7) gives an idea of the importance of the researchers involved and the scope of the topics discussed: R. L. Anderson, Trygve Haavelmo, Harold Hotelling, Leonid Hurwicz, Lawrence R. Klein, Tjalling C. Koopmans, R. Leipnik, Henry B. Mann, Jacob Marschak, H. Rubin, Gerhard Tintner, and Abraham Wald discussed topics such as time series analysis and maximum likelihood estimation and identification problems in simultaneous equation models.

Thus, the January 1945 conference and Cowles Monograph No. 10, which laid out the results of this conference, were crucial to the path that econometric research in macroeconomics has followed since then. In particular, the journal brought forth groundbreaking papers that set the trajectory of research on the conditions for identifying structural coefficients of simultaneous equations, regarding problems of bias when the estimate of simultaneous equations employs least squares methods that are suitable for only one equation, as well as on full and bounded information maximum likelihood methods (Dimand, 2019, p. 8). *The cornerstones of modern empirical research in macroeconomics were laid at that event and in the publication that followed it, especially concerning simultaneous equation models*.

But the role of the Cowles Commission went beyond advances in the use of econometric techniques to produce empirical evidence in macroeconomics. The commission also played a key role in an issue that is of direct interest to this paper: the relationship between theory and empirical research. Although the commission's research had always had some connection with economic theory, the latter was not the direct object of research (Malinvaud, 1983, p. 2). This picture will begin to change with the entry of Oskar Lange and Jacob L. Mosak in 1939: "their Cowles Commission Monographs, respectively Nos. 8 and 7, both published in 1944, *Price flexibility and employment*, and *General-equilibrium theory in international trad*e, were the first ones to deal with formalized economic theory" (Malinvaud, 1983, p. 2).

The second decisive moment in the redefinition of the relationship between theory and empirical macroeconomic research, with the emphasis being progressively shifted to the theoretical foundations of empirical research, would happen through 1942-1943, with the entry of Leonid Hurwicz and Trygve Haavelmo, in addition to the already mentioned participation of Jakob Marschak. Malinvaud (1983, p. 2) explains that the share of theoretical-themed papers jumps from something around one-third of the titles until 1950 to two-thirds of the titles in the following years, which led the committee to change its motto from "science is measurement" to "theory and measurement".

The tools for empirical discussion of macroeconomic policies combined with theoretical explorations in economics were thus established. With the development of econometric techniques for estimating simultaneous equation models and an emphasis on theoretical discussion, the Cowles Commission laid the foundation for what was to become the practice of empirical discussion in macroeconomics from the mid-twentieth century on.

Edmond Malinvaud (1998) describes the rise of econometric models as a tool for empirical assessment of macroeconomic policies in the United States and Europe, starting in the 1950s. Although Jan Tinbergen's main econometric contributions in Europe began in the 1930s (Tinbergen, 1937), in the post-war period, it was after the Klein-Goldberg model in the United States in 1955, and since 1957 in Europe, that econometric models began to gain wide acceptance as tools for empirical assessment of macroeconomic policies (Malinvaud, 1998, p. 330).

This expansion of the use of econometrics for the assessment of macroeconomic (fiscal and monetary) policies was strongly influenced by the wide acceptance of Keynesianism as a fundamental tool of economic management. As Malinvaud (1998, p. 330) explains, the acceptance of Keynesianism was a result of the concern to avoid the return of the interwar crisis, since this theory proposed precisely to combine fiscal and monetary measures, to ensure full employment with price level control.

Therefore, the use of econometric models to assess macroeconomic policies was born associated with Keynesianism, having as its central motivation the search for full employment with inflation control. An important step in this direction was taken by Henri Theil, who, based on his experience in Holland published *Economic forecasts and policy* (Theil, 1958), in which he discussed econometric methods for studying economic policies and predicting their effects.

The growing use of macroeconometric models was accompanied by an optimism that expressed itself in ever more extensive models, surrounded by growing expectations, which materialized in the idea that it would be possible to design an optimal macroeconomic policy based on evidence, *evidence being understood as the results of these econometric models*. This led "economists, armed with their dynamic models, to find themselves in a position even similar to that of engineers called upon to optimally direct the trajectory of a rocket" (Malinvaud, 1998, p. 330).[8]

This optimistic expectation that econometric models could provide empirical evidence for the adoption of an *optimal macroeconomic policy* would be severely shaken in the 1970s. In that period, there was the combined experience of economic

8. "*Des économistes, dotés de leurs modèles dynamiques, se virent même dans une position semblable à celle d'ingénieurs appelés à diriger au mieux la trajectoire d'une fusée*".

stagnation and inflation in the United States, which became known in journalistic terms as *stagflation*.

As will be seen later, the inability of Keynesian models to initially deal with this hitherto unprecedented conjuncture motivated Lucas (1976) critique, the first to theoretically challenge the use of econometric models to assess the scope of macroeconomic policies. This critique was the most far-reaching and motivated the search for structural parameters (deep parameters) that reflect the choices of rational maximizing agents when faced with the possibilities of choice and the very policies adopted, in the approach that became known as rational expectations.

3 LUCAS CRITIQUE

The discussion about the problems with the use of evidence in macroeconomics begins with Lucas critique of Keynesian econometric models (Lucas, 1976), as this critique was one of the first to significantly affect one of the main sources of evidence in the macroeconomic debate: the results of econometric models. At the same time, Lucas critique offers a unique opportunity to study the difficulties of the macroeconomic debate with empirical evidence.

Lucas critique, as presented in *Econometric policy evaluation: a critique* (Lucas, 1976, p. 41), is summarized at the end of the chapter, where it is presented as "a single syllogism": since the structure of an econometric model is constituted by the rules of the agents' optimizing behavior, any policy change will transform the structure of the model, as it changes the relevant data for these agents' decision-making process.

As can be seen from the previous quote, *Lucas critique was fundamentally directed at the use of macroeconomic models in the evaluation of public policies*. It can be said, therefore, that the critique was perhaps the first theoretical effort to question the grounding of public policies in empirical evidence based on the results of econometric models.

Lucas critique claimed that changes in economic policies altered the very way in which these policies affected the economy. The reason is that by being rational – that is, using all available information – agents would anticipate the consequences of new macroeconomic policies and consequently change their behavior. This would have damaging implications for the use of econometric models in formulating and, especially, predicting the effects of economic policies. As Lindé (2001, p. 896) explains, after Lucas critique, past behavior would no longer be a valid reference to estimate the effects of alternative policies, and the parameters of econometric models in reduced form would no longer be constant.

As is well known, the reduced form of an econometric model is built from a structural model – that is, from a model of equations built from theoretically derived relationships. The reduced form is nothing more than an algebraic arrangement, in which the endogenous variables are placed as a function of the exogenous variables. Therefore, it is far less detailed than the structural form. Because they are simpler, the theoretical foundation of models in reduced form is less theoretically demanding.

The importance of Lucas' (1976) critique in the macroeconomic debate cannot be overstated: it has been assimilated by a majority in academia as a fundamental step in the modernization of economic theory, at least as far as its dominant paradigm is concerned, as exemplified by Hall's (1996) assessment of the importance of Lucas's contribution, written in light of his awarding of the 1995 Alfred Nobel Memorial Prize in Economic Sciences. According to Hall (1996, p. 38), the effect of Lucas's critique was to train subsequent generations of economists to develop macroeconomic models in a way that was rigorously consistent with microeconomic fundamentals, which would have affected not only the field of applied economics but also economic theory.

As Lucas' (1976) very critique concerns how evidence is used in the macroeconomic debate and how the use of that evidence affects the accuracy with which forecasts are made, the debate concerning this critique has involved from the beginning a *discussion about the ability to make forecasts about the behavior of the major economic aggregates*. In actuality, Lucas (1976) used the Phillips curve and its inverse relationship between unemployment and inflation, one of the basic instruments of active macroeconomic policy, as an example of his argument.

It is curious to notice that the Phillips curve *is one of the few cases of empirical observation giving rise to the production of a new theory*, something that should be common if the production of economic theory was usually based on evidence.[9] However, according to Lucas (1976, p. 40), the Phillips curve failed to anticipate the so-called *stagflation* of the 1970s in the United States. According to the author,

9. As is known, the origin of the Philips curve is his paper *The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861-1957* (Phillips, 1958). Even though evidence can be found in the literature of some passages where earlier authors have identified some inverse relationship between unemployment and inflation, possibly, according to Humphrey (1985), going back to John Law (1621-1729), it is generally accepted (Gordon, 2011) that the link between unemployment and inflation was formally established by Phillips' estimated regression, represented by: $w_t = -0{,}90 + 9{,}638U^{-1{,}934}$ (Phillips, 1958, p. 290), in which w_t is the annual rate of change of nominal wages in a percentage and *U* the unemployment rate.

This empirical identification produced a wave of theoretical innovation, although sometimes rendering the relationship ineffective, as in the case of the version of the Phillips curve with rational expectations, which nullifies the trade-off between inflation and unemployment even in the short run. To discuss the revisions of this curve since its inception would be impossible within the scope of this paper. The interested reader is advised to refer to, among several possible references, Gordon (2011).

there would then have been instability in the parameters of the Phillips curve, caused by the reaction of rational agents to the macroeconomic policies of the period.

Therefore, much of the strength of Lucas' (1976) critique stems both from its presentation as a logical – and therefore in principle irrefutable – syllogism and from the apparent inability of the Keynesian models of the time – in particular, the Phillips curve – to explain and predict the combination of economic stagnation and inflationary acceleration of the 1970s. Proper consideration of Lucas's critique, therefore, requires that these two aspects of his triumph be considered.

Lucas critique, if taken *superficially*, is really a syllogism. If agents alter their behavior in response to changes in economic policy, reduced-form models that do not incorporate agents' reactions to policy changes are doomed to irrelevance at best and to produce erroneous forecasts at worst.

This superficial reading does not, however, exhaust the issues associated with Lucas' (1976) critique. In fact, this critique involves at least two other questions (one theoretical and one empirical), and the two questions do not boil down to simple syllogisms.

The theoretical question concerns the type of behavior of agents when faced with changes in economic policy so long it is accepted that they respond rationally to policy changes.

This question, as is well known, was answered by Lucas and others with the rational expectations model: agents are rational maximizers, in the sense that they have and use all the available information and thus correctly anticipate the consequences of macroeconomic policies, largely cancelling out the effect of any discretionary policies. The solution would be to establish clear rules that influence agents' expectations, as is the case with the inflation targeting regime, rather than to adopt discretionary macroeconomic policies, such as an active fiscal policy. Obviously, the assumption of rational expectations cannot be considered a syllogism, but a hypothesis to be empirically assessed.

However, when considering its consequences in terms of models of economic policy analysis, Lucas's critique is surrounded by paradoxes, especially the one related to one of its recent and important developments, which is the dynamic stochastic general equilibrium models (DSGE).[10] These models are considered mainstream in macroeconomics, particularly in monetary policymaking. This type of model, of course, also enjoys great popularity in Brazil, as indicated by the works of Vereda and Cavalcanti (2010), Cavalcanti and Vereda (2011), Ferreira (2015), Areosa and

10. Sergi (2018, p. 2) identifies the DSGE models as part of the new neoclassical synthesis that seeks to respond to Lucas critique. See also Hurtado (2013; 2014).

Coelho (2015), and Nunes and Portugal (2018), just to mention some of the most representative works in the country.

Nevertheless, despite being an offshoot of Lucas critique, DSGE models, curiously, do not meet the criteria for parameter invariance. Even their main advocates, such as Fernández-Villaverde and Rubio-Ramírez (2008, p. 84), claim that it is difficult to accept that the parameters of DSGE models are structural. For example, most of these models specify a stable production function, with constant elasticity of output with respect to capital – something unacceptable, considering Lucas' (1976) criticism, since changes in relative factor prices would induce the development of new technologies. Also according to the authors, problems of this type would affect almost all dimensions of a modern DSGE model.

In fact, the adoption of the rational expectations hypothesis is no guarantee of the stability of model parameters, as demonstrated by Estrella and Fuhrer (1999). By testing models with optimizing behavior based on rational expectations versus simpler models without the same behavior, the authors observed that there is little evidence that backward-looking models are unstable, unlike models with rational expectations, which show *clear evidence of instability* (Estrella and Fuhrer, 1999).

Despite their instability, it is still possible to defend rational expectations models if i) the instability of the parameters is also significant in the Keynesian models used so far; and ii) the problems with the Phillips curve and other Keynesian models used in the 1970s are *actually* a consequence of changes in agents' behavior in response to changes in macroeconomic policies.

If other factors have also played a role in addition to, or as a substitute for, eventual changes in an agents' behavior, a subsequent question involves evaluating whether these other factors could be incorporated coherently into the Keynesian models then in force, or whether they could only be included in those models with the addition of ad hoc hypotheses – that is, hypotheses elaborated from the very fact that is intended to be explained, to avoid discrediting the theory.

According to Goutsmedt et al. (2019, p. 535), Lucas' (1976) critique was widely accepted and incorporated into the dominant paradigm by most economists as an unquestionable principle that immediately disqualified Keynesian models, ignoring not only the issues mentioned above but also a whole series of empirical evidence that was presented by Keynesian economists in the debate about Lucas critique during the 1970s and 1980s.

The debate about Keynesian models that followed Lucas' (1976) critique clearly illustrates the difficulties in assessing evidence when it comes to discussing macroeconomic policies, as well as the overvaluation of theory to the detriment of empirical evidence. To this end, it is necessary to emphasize in Keynesian answers

what concerns *only* the effect that agents' expectations may have on the structure of the model, disregarding specific questions about the rational expectations hypothesis. This is because the debate over Lucas (1976) often combines the empirical question with the theoretical discussion about the rational expectations hypothesis.

Indeed, as Goutsmedt et al. (2019) explain, the core of the Keynesian critique of Lucas (1976) lay in its *practical relevance*. Malinvaud corroborates this point, clarifying that the empirical evidence for the validity of Lucas critique remained very limited, even more than twenty years later: the "little illustrative models" (*les petits modèles illustratifs*) presented by Lucas and other authors would only demonstrate a possibility, and have not been subject to more accurate empirical testing (Malinvaud, 1997, p. 21). On the same point, Malinvaud would write a year later something that would become the central argument to explain the empirical irrelevance of Lucas' (1976) critique: that private agents do not care about monetary and budgetary decisions unless they affect them directly (Malinvaud, 1998, p. 335).

Hence, it is reasonable to inquire from where the success that Lucas' (1976) critique enjoyed in academia originated, almost immediately upon its release. It seems that this success resulted not only from its strength as a "simple syllogism" – which was seen not to be so simple – but also from empirical evidence supporting the critique. However, this evidence was not direct: it was related to the apparent inability of Keynesian models to anticipate economic fluctuations – particularly the combination of high unemployment and inflation of the 1970s. This inability was taken by critics of Keynesianism as evidence of the relevance of Lucas critique.

In fact, at the empirical level, Lucas' (1976) critique has not been directly tested systematically and repeatedly. What was empirically tested were some of its hypotheses and predictions, *based on the rational expectations model*. Thus, there was no dissociation of the critique from the theoretical school it contributed to the foundation of. Some of the classic works in this regard are those by Lucas (1973), Sargent and Wallace (1973), Sargent (1976), and Barro (1977).

This predictably led to the Keynesians' response involving empirically assessing the existence of structural breaks and parameter instability in their models following a change in economic policy, with an emphasis on empirical analysis of the Phillips curve, the central macroeconomic policy tool, according to this generation of Keynesians. This occurred despite Lucas' (1976) critique being much broader, concerning the structural stability of models in the face of the possibility of changes in expectations in response to changes in macroeconomic policies (Goutsmedt et al., 2019).

The first Keynesian response was given by Blinder in his book *Economic policy and the great stagflation* (Goutsmedt et al., 2019, p. 10). Other authors fol-

lowed, such as Otto Eckstein, with his books *The great recession, with a postscript on stagflation*, and *The DRI model of the U.S. economy*, Lawrence R. Klein (1985) and Robert J. Gordon (1975; 1984; 2011). All of these responses emphasized that the conventional Keynesian model and the Phillips curve in particular adequately described the stagflation situation of the 1970s, provided that they incorporated the supply shocks in energy and agricultural products of the period.

However, these responses have simply been ignored, which has led Keynesians like Blinder and Malinvaud to consider New Classical macroeconomics an illegitimate "palace coup" since it is devoid of an empirical basis (Goutsmedt et al., 2019, p. 22). It is important to stress here that one is not discussing the technical quality of these responses, only the fact that they have been ignored, and Lucas' (1976) critique has been considered a milestone in the use of evidence in macroeconomics, totally disregarding its questioning from an *empirical* point of view by renowned professionals.

Even more surprisingly, Goutsmedt et al. (2016, p. 11) point out that Lucas's critique was *nothing new* – in fact, something that was acknowledged in passing in a short footnote by Lucas (1976, p. 20) himself. The exact same idea was explicitly discussed by Jacob Marschak and Jan Tinbergen but was dropped in the following decades. The interesting question would then be to ask what would have allowed Lucas (1976) to reopen the discussion (Goutsmedt et al., 2016, p. 13).

There are thus two important omissions, from an empirical point of view, regarding Lucas' (1976) critique: the macroeconometric papers that anticipated and discussed the problem of parameter stability; and the empirical responses that questioned the relevance of the critique for Keynesian models. It proves very difficult to frame these omissions from the perspective of an evidence-based subject. A scientist normally considers the evidence that preceded his work, as well as any empirical evidence that questions his result.

These facts indicate that there is indeed a significant problem in the relationship between macroeconomics and empirical evidence. Lawrence Summers (1991) argues along these lines, and his critique will be considered below.

4 SUMMERS AND THE "SCIENTIFIC ILLUSION" IN EMPIRICAL MACROECONOMICS

The most recent critique of the way evidence is used in macroeconomics through econometric models is that of Lawrence Summers (1991), in his paper *The scientific illusion in empirical macroeconomics*. This criticism is relevant because he is an economist with extensive experience in macroeconomic policymaking, having worked in the United States Treasury Department and the World Bank, and having served on the National Economic Council during the first administration of former United States president Barack Obama.

Thus, Summers (1991, p. 144) defines what he calls "the scientific illusion in empirical macroeconomics", namely the beliefs that: i) empirical efforts in macroeconomics should focus on the deep structural parameters related to preferences and technologies; ii) empirical work in macroeconomics should test hypotheses rigorously derived from theory; and iii) sophisticated statistical techniques are important to distinguish causal relations in systems with many interdependent variables. As seen in the previous section, the aforementioned beliefs are direct results of Lucas critique. These beliefs, according to Summers (1991, p. 144) "form the core of what I consider the scientific illusion in empirical macroeconomics".

Summers' (1991) critique of the role of evidence in the macroeconomic debate can be summarized as follows: macroeconomic theory is divorced from empirical observation and overemphasized by the "failure of empirical work to deliver facts in a form where they can be apprehended by theory" (Summers, 1991, p. 144).

Thus, Summers (1991) identifies two fundamental difficulties in the relationship between macroeconomics and empirical evidence. One concerns the excessive weight of theory in the debate. The other concerns the type of evidence that is produced. The two problems are two sides of the same coin.

Therefore, Summers (1991), in criticizing the supremacy of theory over empirical evidence in the macroeconomic debate, is not an isolated case. The excessive weight of theory is also identified and criticized by Juselius (2010, p. 2), who calls for greater prominence of empirical analysis over theory, not only to provide more solid grounds for analysis but also to underpin new theories.

It would be reasonable to expect that empirical evidence would not only stimulate the production of new theories, as Juselius (2010) calls for, but also serve as a basis for predictions that can be empirically tested. However, this is not the case, as Summers (1991, p. 144) points out, because the vast majority of theoretical macroeconomics, despite emphatic statements in favor of rigor and generality, neither starts from empirical observations nor results in empirically verifiable predictions.

In fact, Summers (1991, p. 131-132) sets the crucial role of empirical evidence in the development of new scientific theories in opposition to the almost irrelevance of such evidence for the development of new economic theories. He mentions, as an example of the crucial contribution of empirical evidence to stimulate the development of new scientific theories, the role of the Hubble telescope in the creation of the Big Bang theory; or the study of fossils, fruit flies, and the DNA of various species for the most modern formulations of the theory of evolution.

To stack up the theoretical work of economists against those of scientists who develop new theories from empirical evidence, Summers (1991) lists several

works that have helped shape current macroeconomics, with few or no references to econometric models, such as *Models of business cycles*, by Robert Lucas (1987); *Dynamic macroeconomic theory*, by Thomas Sargent (2009); *Growth theory: an exposition*, by Robert Solow (1970); and *Asset accumulation and economic activity*, by James Tobin (1982).

Since the evidence provided by econometric models – especially those that incorporate more sophisticated techniques – is weak and problematic, the theory behind the models ends up being more important than the empirical evidence itself. Also, according to Summers (1991), contrary to the practice in the natural sciences, replications of econometric results to test estimated parameter values are rare. For Summers (1991, p. 133), this is a consequence of the fact that "the results are rarely an important input to theory creation or the evolution of professional opinion more generally".

Effectively replicated econometric work, according to Summers (1991, p. 133), usually involves qualitative aspects, rather than estimating structural parameters or testing a hypothesis. The reason for this – despite the significant methodological innovations involved in Thomas Sargent's deep parameter approach or Christopher Sims' vector autoregression (VARs) approach – would be that researchers in both lines mistake methodological advances for advances in substance (Summers, 1991, p. 134).

Summers' criticism of Sargent's approach to the estimation of deep parameters – that is, the structural parameters that describe the fundamental behavior of consumers and firms[11] – is focused on two papers by Hansen and Singleton (1982; 1983). These two papers are recognized for their econometric value, as shown by the Frisch medal awarded to Lars Peter Hansen and Kenneth Singleton in 1984. Given the outstanding character of these econometric papers, their basic flaws represent this entire genre of work (Summers, 1991, p. 134).

The first important issue to be stressed is that even if the hypothesis under test is not rejected, Summers (1991) still points out that the structural parameters estimated in models such as those of Hansen and Singleton (1982; 1983) would hardly be taken seriously. In addition to limitations in data access and use, *it is uncommon for estimated structural parameters to be used to make predictions regarding the effects of macroeconomic policies*. As Summers notes, even though Hansen and Singleton (1982; 1983) have estimated the structural parameters of the utility function of representative consumers, as recommended by Lucas' (1976) critique,

11. As explained by Low and Meghir (2017, p. 35), fully specified structural models adopt explicit assumptions about the goals of economic actors, their environment, and their information set, also specifying the choices that can be made; thus, allowing the individual optimization problem to be solved as a function of the information set.

it is unlikely that anyone would use these estimates to calculate the effects of a tax cut (Summers, 1991, p. 136).

Here, we have another key point of this work. Given the large volume of econometric papers in specialized journals, why are the estimated values of the structural parameters hardly ever used in policy simulations, unlike the parameters estimated in laboratory experiments, which are subsequently employed in engineering applications?[12] It is the fact, noted by Summers (1991), that these structural parameters are usually ignored when discussing the effects of public policies.

The answer to this question, both regarding Hansen and Singleton (1982; 1983) and any other model concerned with estimating structural parameters, is simple: the heroic assumptions and frequent problems in specifying the structure of models – even more so in the case of those employing the concept of the representative consumer – generate so many uncertainties that they make it impractical to apply the parameter estimated in this way to evaluate public policies, even if the model does not reject the assumptions under test.

It should not come as a surprise that Summers (1991, p. 137) considers that "Hansen and Singleton's work creates an art form for others to admire and emulate but provides us with little new knowledge". An equally severe judgment would apply to any similar attempt "to test a highly restricted and surely incorrect structure using elaborate methods which do not shed light on the cause of any deviations of data from theory" Summers (1991, p. 137).

According to Summers (1991), the main goal of Hansen and Singleton (1982; 1983) was to test the relationship between consumption and asset prices, based on a representative consumer model with rational expectations. Summers (1991, p. 135) then identifies some relevant general problems in Hansen and Singleton's approach. These authors offer no indication as to the origin of their failure, whether in the logic of the theory itself or in the auxiliary hypotheses made to test that theory.

This is an important argument: the more complex the theoretical models and the more sophisticated the estimation techniques to try to empirically validate these models, the more auxiliary assumptions have to be made to obtain any result, and, as a consequence, it becomes increasingly difficult to discern whether the result (the rejection or non-rejection of a given hypothesis) is a consequence of the hypothesis itself or of the additional requirements derived from the estimation technique that is needed to empirically evaluate the hypothesis.

12. It is important to emphasize that this is not about questioning the *theoretical* bases of these models; for example, the hypothesis of the representative consumer, given the obvious heterogeneity of consumers. The discussion in this paper is restricted to the question of the use of empirical evidence in macroeconomics, accepting the models that are seen as mainstream.

In fact, several problems are involved in estimating a theoretical model with micro-foundations that cannot be directly observed and that require complex statistical techniques for its empirical evaluation. As Hendry and Muellbauer (2018, p. 304) point out, first, even if there is only one theory to be empirically evaluated – that is, even if the *theoretical consistency* condition is met – this does not mean that there is only one model for empirical estimation: there are different ways to implement a model empirically, varied ways to consider variables that cannot be observed and different ways to handle expectations.

Nor is there just one way to ensure empirical consistency. There are different ways of measuring data; data revisions happen; it is not always possible to calculate precisely the uncertainty of estimates; and the selection criteria for the different empirical models are not unique and homogeneous. All these issues generate an inevitable uncertainty regarding the values of the estimated parameters. As a consequence, it is not surprising that these parameters hardly contribute to the estimation of macroeconomic policies.

Second, if the hypothesis that drives the study is rejected due to details in the implementation of the empirical test, this rejection provides no indication of the type of change needed in the theory. Even more seriously, the rejection of a theoretical hypothesis because of the particular implementation characteristics of an empirical test provides no stimulus for further theoretical developments with respect to the hypotheses being tested (Summers, 1991, p. 135).

As for the alternative technique of examining empirical evidence in vector autoregression macroeconomics, Summers' (1991) verdict is no more favorable. He endorses what he calls Sims' (1980a) "destructive argument" against structural models (Summers, 1991, p. 137).[13] But Summers (1991) is less condescending about the possibility of VAR-based modeling providing input for the formulation of macroeconomic theories and policies.

In fact, VAR-based models are usually subject to severe criticism. As Fabio Canova (1995, p. 57) explains, critics point out that the methodology of VAR models bears little relation to economic theory, is based on a set of unsustainable assumptions, and is essentially flawed, being subject to Lucas critique. It is thus a paradox that after the academic success of this critique, VAR-based models have achieved popularity.

Indeed, Summers' (1991) critique uses Ben Bernanke's (1986) paper, *Alternative explanations of the money-income correlation*, as a paradigmatic example of the problems of empirical evidence generated by this type of model. According to

13. Summers mentions only Sims (1980a) as a fundamental reference for VAR-based modeling, but Canova (1995) also mentions Sims (1972; 1980a; 1980b) and Sims, Goldfeld and Sachs (1982).

Summers, the only valid conclusion from VAR-based models such as Bernanke (1986) is that "the only firm conclusion reached is that structural interpretations of VARs are very sensitive to the model one assumes" (Summers, 1991, p. 138). Such a rigorous conclusion should come as a surprise since it is a feature of VAR-based models that they do not require a theoretical background, only that they select the variables to be included and define the lag structure.

Given the lack of theoretical substance in these models, estimates and forecasts based on VARs are strongly affected by the inclusion or exclusion of variables (Braun and Mittnik, 1985) and their gap structure (Hafer and Sheehan, 1989). In the context of this paper, it is particularly relevant to note that the structure of the gaps also severely affects the economic recommendations suggested by the models (Hafer and Sheehan, 1991).

Here, we reach another point of fundamental importance in this paper. On the one hand, models that seek to estimate structural parameters impose an overly demanding theoretical framework on the data, which makes the results lack robustness. As a result, much is invested in statistics to evaluate the conformity of the model with the data, an unavoidable necessity due to the overly restrictive nature of the hypotheses that are adopted – such as the hypothesis of a representative consumer (Summers, 1991, p. 136-137).

On the other hand, forgoing theoretical support in the data analysis – as VAR-based models do – not only gives room for a large number of alternative theoretical hypotheses to explain the same result but also makes this type of model very susceptible to arbitrary selections by the researcher, regarding both variables and gaps. This renders this type of modeling of little interest for the design and evaluation of macroeconomic policies.

In fact, there seems to be a problem in the relationship between macroeconomics and empirical evidence that seems to oscillate between two extremes: either theory "rules" the evidence, as in the case of structural models, or statistical relations between historical series start to "suggest" a link between them (VAR-based models), and theory becomes something secondary, at the discretion of the researcher.

Next, an even more radical criticism of structural models will be addressed, in which this type of model is considered a step *backwards* concerning the development of the scientific character of macroeconomics, despite appearances to the contrary.

5 ROMER AND THE LOSS OF MACROECONOMIC'S SCIENTIFIC CHARACTER

Romer's (2016) critique is even more radical and poignant than Summers' (1991). Given the arsenal of mathematically complex theoretical models and the wide variety and availability of statistical techniques and series, it is not surprising that

Romer's quote, which serves as the epigraph to this paper, is somewhat shocking when he states that macroeconomics has been losing its scientific character. One should also note the academic relevance of the author of such a severe diagnosis: Romer was awarded the 2018 Alfred Nobel Memorial Prize in Economic Sciences, alongside William Nordhaus.

In his work, *The trouble with macroeconomics*, Romer (2016, p. 9) blames the dominant theoretical paradigm in macroeconomics for rejecting empirical evidence that contradicts its theoretical postulates, in favor of the authority of economists such as Robert Lucas. According to this author, this refusal would be aggressive towards any criticism, *even if the criticism is based on evidence*. More seriously, Romer (2016) blames the so-called *mainstream macroeconomics* for promoting retrocession in this theoretical field, especially with its habit of *deliberately ignoring evidence*. This would be the case, for example, when mainstream authors claim that the Central Bank has no way of affecting the interest rate that functions as the basis of the economic system and that monetary policy would be irrelevant.

Mentioning deflation early in Paul Volcker's tenure at the United States central bank, the Federal Reserve (FED), achieved from sharply contractionary monetary policy, Romer (2016, p. 4) notes that if the FED can cause the interest rate to change by 500 basis points, it seems absurd to inquire whether monetary policy is relevant. Romer (2016, p. 10) suggests that a statement that so "blatantly" contradicts established facts suggests a lack of interest by economists in the very facts.

Romer's (2016) very piece offers an example of the macroeconomic field's difficulties in dealing with empirical evidence. This author presents statistical series that show how the United States central bank caused deflation at the beginning of Paul Volcker's administration – between 1979 and 1983 – by raising the real interest rate on federal funds, which also resulted in a significant reduction in economic activity and increased unemployment. To these statistics, Romer (2016, p. 2) counters the work of Jesús Fernández-Villaverde (2010), defined by him as one of the leading mainstream economists, and who states that, despite the "impressive empirical case" of those who believe that money affects economic fluctuations, he is not fully convinced of the importance of money, "outside the case of large inflations".

The empirical evidence presented by Romer (2016), his criticism, and the skepticism expressed by Fernández-Villaverde (2010) definitely cannot coexist in the same research field without a sense of embarrassment. The effective use of evidence should limit the debate, at first, only to issues of secondary importance

in the macroeconomic field, helping to pacify fundamental questions.[14] Why is the debate still going on if there is a variety of statistics available to be used as evidence?

This question raises another important issue: it is not enough to have evidence, even if in the form of historical series built with statistical rigor and precision – which is often not the case. *It is necessary to define what can validly be considered as legitimate evidence to support a statement in macroeconomics*, *and how this evidence can be properly used in the debate*. As seen earlier, different econometric models can account for the same theory – that is, there are different ways to estimate a theoretical hypothesis empirically, there are several data alternatives to measure the same variable, there is the problem of variables that cannot be observed, and there are a variety of equally valid criteria for selecting an empirical model.

These factors aggravate the embarrassment resulting from the lack of generally accepted protocols on the scientifically adequate way to use these statistical techniques in order to support a theoretical proposition. Without such protocols, the macroeconomic debate that should define what fiscal and monetary policies are appropriate for a given purpose ends up relying on other factors, such as academic authority, the mathematical sophistication of theoretical models, or both.

Thus, given the profusion of statistical and econometric methods whose criteria are subject to strong questioning, even the most basic theoretical relations that guide macroeconomic policy proposals are immersed in bitter controversies, in which the academic prestige of the participants and the complexity of the theoretical models presented seem to have greater importance than the available evidence. In Romer's words (2016, p. 8), "progress in the field is judged by the purity of its mathematical theories, as determined by the [academic] authorities".

Indeed, even an empiricist author like Reiss (2008, p. 2) is forced to recognize that, as far as economics is concerned, even a basic concept such as a *firm* cannot be directly observed, being something of a completely different nature from a collection of machines, buildings, CEOs etc. Reiss' (2008, p. 3) recommendation, therefore, is to *acknowledge that there is a plurality of methods to construct and gather evidence*, which involves, in addition to direct observation of evidence from the senses, also statistical methods, such as number-indexes; econometric regressions; analysis of variance (Anova); mathematical modeling; computer simulation models; experimental economics; mental experiments; expert opinion etc. Therefore, the possibility of evidence-based fiscal and monetary policies depends rather on an appreciation of the various types of evidence in the debate concerning basic macroeconomic fundamentals.

14. Except in extraordinary moments, what Thomas Kuhn (1998) called *scientific revolutions*.

Any naïve empiricist perspective that minimizes the complexity in the construction of evidence in macroeconomics runs into the fact that the econometric evidence widely used in the debate is built on techniques that offer many degrees of freedom to the analyst, perhaps even excessive degrees of freedom, which may allow for lax manipulations of the results, especially when there are no generally accepted protocols as to the appropriate level of use of these degrees of freedom. These problems are first illustrated by Romer (2016) with the well-known identification problem. In this regard, criticizing economists who lack the commitment to historical evidence, he notes that modeling strategies that allow for more variables and more *imaginary shocks* provide additional degrees of freedom since more variables aggravate the identification problem (Romer, 2016, p. 10).

Romer (2016, p. 11-12) discusses the identification problem with a simple log-linear labor supply and demand model with random shocks. To predict the effect of a policy change, one needs to estimate the elasticity of labor demand. As is well known, in this case, the solution to the identification problem is to impose some constraint on the supply curve. Romer (2016, p. 10-11) imposes two alternative restrictions: one in which the supply curve is perfectly inelastic; and another in which it has a positive shift and passes through the origin. Each constraint produces very different results for the demand curve. Therefore, without additional information, no result has any meaning. Herein lies the possibility of arbitrary manipulations, according to the preferences of the researcher.

Romer (2016) also provides other examples of econometric model estimation that allow too much freedom to the analyst's arbitrariness. He mentions, for example, the model of Smets and Wouters (2007), which has seven variables and 49 parameters to estimate, with only seven equations, so 42 of the 49 parameters have to be estimated with information *other* than the time series of x (Romer, 2016, p. 12). The inclusion of rational expectations further exacerbates the identification problem by increasing the number of parameters, which have to be specified based on information other than the time series of the independent variables (Romer, 2016, p. 12).

Because of this excessive freedom given to the arbitrariness of analysts, the solutions found have become increasingly opaque: "a discussant or referee cannot say that an identification assumption is not credible if they cannot figure out what it is and are too embarrassed to ask" (Romer, 2016, p. 15). Such opacity in the use of the degrees of freedom provided by the models has increased over time. Romer (2016) goes on to say that assumptions about the distribution of error terms are a good place to "bury things", simply because "hardly anyone pays attention to them" (Romer, 2016, p. 15).

There are also problems when using Bayesian methods. Back to the example of estimating the elasticity of the demand curve for labor, in the case of using a

Bayesian method, Romer (2016, p. 15) notes that by manipulating the backgrounds for the supply curve, it is possible to change the subsequent outcomes for the elasticity of demand, until you get the one desired.[15]

In face of all these issues, Romer (2016, p. 19) concludes that

> perhaps this time, macroeconomists should admit that the wreckage runs so deep that they should abandon the quest for the sacred simultaneous equation model. It might be wiser to adopt the messy methods that medical researchers have used to make discoveries that were implemented and actually improved health.

That is, rather than dogmatic fidelity to a particular method of empirical analysis – such as structural models – it may be more productive to consider a wider range of evidence of different natures, even if the resulting method is not formally elegant. The different pieces of evidence may reinforce each other or, alternatively, provide different results and thereby broadening the perspective of the analysis, possibly even stimulating new theoretical developments. This point will be taken up again in the conclusion.

6 BY WAY OF CONCLUSION: IS MORE EVIDENCE-BASED MACROECONOMICS POSSIBLE?

Reiss (2008, p. 162) warned that econometric models are almost always poorly specified and that, in socioeconomic analysis, structural breaks are frequent. Consequently, a sophisticated model with complex techniques involving the estimation of structural parameters may end up being of less use than a simpler model when it is necessary to make macroeconomic forecasts, especially those related to policy choice: indeed, more complex models often fail to demonstrate their superiority over simpler models (Reiss, 2008, p. 162). Certainly, the problems with macroeconometric models, in general, are so numerous and with such diverse sources that David F. Hendry (2018, p. 19) inquires what would be a possible criterion for selecting among them, concluding that, historically, the criterion has been conformity with mainstream accepted macroeconomic theory – that is, it has been *internal credibility*, rather than verisimilitude.

This greater weight of theory often ends up translating, in practice, into the greater importance of academic prestige when deciding debates about macroeconomic theory and policy, as charged by Romer (2016). Therefore, there is no shortage of statistical and econometric techniques to address empirical evidence in an attempt to develop evidence-based macroeconomic policies. Instead, the problem

15. Romer (2016, p. 15) quotes several authors to support his criticism of Bayesian methods in producing empirical evidence in macroeconomics, especially regarding the Smets and Wooters (2007) model; among them, Iskrev (2010) and Komunjer and Ng (2011), who show that without background information, the Smets and Wooters (2007) model is not identified.

seems to lie in the way in which these techniques are used, and the excessive importance attached to theoretical models at the expense of empirical analysis.

Critical awareness must be developed toward this type of researcher bias, demanding that all procedures of empirical analysis be made explicit and critically examined by peers, regardless of the researcher's theoretical attachment and fidelity to the principles of a school. The use of different econometric models and empirical evidence – using different data by researchers belonging to different theoretical schools – may to some extent generate results that are not controlled by any one researcher, or group of researchers, even more so if they are combined with other techniques, such as simulation models, or even the analysis of simple descriptive statistics. This can play a role similar to that of randomized clinical trials, even if it does not involve the same degree of randomization that characterizes such trials.

These considerations suggest that efforts toward building more evidence-based macroeconomics should primarily include the simultaneous consideration of different types of evidence and openness to debate in macroeconomics.

REFERENCES

AREOSA, W. D.; COELHO, C. A. Utilizando um modelo DSGE para avaliar os efeitos macroeconômicos dos recolhimentos compulsórios no Brasil. **Pesquisa e Planejamento Econômico**, Brasilia, v. 45, n. 3, p. 407-435, Dec. 2015.

BARRO, R. J. Unanticipated money growth and unemployment in the United States. **The American Economic Review**, v. 67, n. 2, p. 101-115, 1977.

BERNANKE, B. S. **Alternative explanations of the money-income correlation**. Cambridge, United States: NBER, 1986. (Working Paper, n. 1842).

BRAUN, P. A.; MITTNIK, S. **Structural analysis with vector autoregressive models**: some experimental evidence. St. Louis: WashU, 1985. (Department of Economics Working Paper, n. 84).

CANOVA, F. The economics of VAR models. In: HOOVER, K. D. (Ed.). **Macroeconometrics**: developments, tensions and prospects. New York: Springer, 1995. p. 57-106.

CAVALCANTI, M. A. F. H.; VEREDA, L. **Propriedades dinâmicas de um modelo DSGE com parametrizações alternativas para o Brasil**. Rio de Janeiro: Ipea, Mar. 2011. (Texto para Discussão, n. 1588).

CHRIST, C. F. **History of the Cowles Commission 1932-1952**. New Haven: Cowles Foundation for Research in Economics/The University of Chicago, 1952.

COWLES COMMISSION FOR ECONOMIC RESEARCH. **Report for 1945**. Chicago: University of Chicago, 1950. Retrieved Mar. 26, 2024, from: https://ia802907.us.archive.org/28/items/in.ernet.dli.2015.132628/2015.132628.Cowles-Commission-For-Research-In-Economics-Monograph-No10_text.pdf.

DIMAND, R. W. **Macroeconomic dynamics at the Cowles Commission from the 1930s to the 1950s**. New Haven: Cowles Foundation for Research in Economics, 2019. (Discussion Paper, n. 2195).

ESTRELLA, A.; FUHRER, J. C. **Are "deep" parameters stable?** The Lucas critique as an empirical hypothesis. Boston: Federal Reserve Bank of Boston, 1999. (Research Department Working Papers, n. 99-4).

FERNÁNDEZ-VILLAVERDE, J. The econometrics of DSGE models. **SERIEs**, v. 1, n. 1-2, p. 3-49, 2010.

FERNÁNDEZ-VILLAVERDE, J.; RUBIO-RAMÍREZ, J. F. How structural are structural parameters? In: ACEMOGLU, D.; ROGOFF, K.; WOODFORD, M. **NBER macroeconomics annual**. Chicago: University of Chicago Press, 2008. v. 22, p. 83-167.

FERREIRA, D. Regra de Taylor e política monetária no Brasil: considerações empíricas a partir de um modelo DSGE para uma pequena economia aberta. **Teoria e Evidência Econômica**, v. 21, n. 44, p. 9-35, 2015.

GORDON, R. J. Alternative responses of policy to external supply shocks. **Brookings Papers on Economic Activity**, v. 1975, n. 1, p. 183-206, 1975.

GORDON, R. J. **Supply shocks and monetary policy revisited**. Cambridge, United States: NBER, 1984. (Working Paper, n. 1301).

GORDON, R. J. The history of the Phillips curve: consensus and bifurcation. **Economica**, v. 78, n. 309, p. 10-50, 2011.

GOUTSMEDT, A. et al. **Criticizing the Lucas critique**: macroeconometricians' response to Robert Lucas. Paris: CES, 2016. (Working Papers, n. 2015.59).

GOUTSMEDT, A. et al. Reacting to the Lucas critique: the Keynesians' replies. **History of Political Economy**, v. 51, n. 3, p. 535-556, 2019.

HAFER, R. W.; SHEEHAN, R. G. The sensitivity of VAR forecasts to alternative lag structures. **International Journal of Forecasting**, v. 5, n. 3, p. 399-408, 1989.

HAFER, R. W.; SHEEHAN, R. G. Policy inference using VAR models. **Economic Inquiry**, v. 29, n. 1, p. 44-52, 1991.

HALL, R. E. Robert Lucas, recipient of the 1995 Nobel Memorial Prize in Economics. **The Scandinavian Journal of Economics**, v. 98, n. 1, p. 33-48, 1996.

HANSEN, L. P.; SINGLETON, K. J. Generalized instrumental variables estimation of nonlinear rational expectations models. **Econometrica**, v. 50, n. 5, p. 1269-1286, 1982.

HANSEN, L. P.; SINGLETON, K. J. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. **Journal of Political Economy**, v. 91, n. 2, p. 249-265, 1983.

HENDRY, D. F. Deciding between alternative approaches in macroeconomics. **International Journal of Forecasting**, v. 34, n. 1, p. 119-135, 2018.

HENDRY, D. F.; MUELLBAUER, J. N. J. The future of macroeconomics: macro theory and models at the Bank of England. **Oxford Review of Economic Policy**, v. 34, n. 1-2, p. 287-328, 2018.

HUMPHREY, T. M. The early history of the Phillips curve. **Economic Review**, v. 71, p. 17-24, 1985.

HURTADO, S. **DSGE models and the Lucas critique**. Madrid: Banco de España, 2013. (Documentos de Trabajo, n. 1310). Retrieved Apr. 10, 2021, from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2307888.

HURTADO, S. DSGE models and the Lucas critique. **Economic Modelling**, v. 44, s. 1, p. 12-19, 2014.

ISKREV, N. Local identification in DSGE models. **Journal of Monetary Economics**, v. 57, n. 2, p. 189-202, 2010.

JUSELIUS, K. **On the role of theory and evidence in macroeconomics**. Copenhagen: University of Copenhagen, 2010. (Discussion Papers, n. 10-12).

KLEIN, L. R. Did mainstream econometric models fail to anticipate the inflationary surge? In: FEIWEL, G. R. (Ed.). **Issues in contemporary macroeconomics and distribution**. London: Palgrave Macmillan, 1985. p. 289-296.

KOMUNJER, I.; NG, S. Dynamic identification of dynamic stochastic general equilibrium models. **Econometrica**, v. 79, n. 6, p. 1995-2032, 2011.

KUHN, T. S. **A estrutura das revoluções científicas**. 5th ed. São Paulo: Perspectiva, 1998.

LINDÉ, J. Testing for the Lucas critique: a quantitative investigation. **The American Economic Review**, v. 91, n. 4, p. 986-1005, 2001.

LOW, H.; MEGHIR, C. The use of structural models in econometrics. **Journal of Economic Perspectives**, v. 31, n. 2, p. 33-58, 2017.

LUCAS, R. E. Some international evidence on output-inflation tradeoffs. **The American Economic Review**, v. 63, n. 3, p. 326-334, 1973.

LUCAS, R. E. Econometric policy evaluation: a critique. **Carnegie-Rochester Conference Series on Public Policy**, v. 1, p. 19-46, 1976.

LUCAS, R. E. **Models of business cycles**. Oxford: Basil Blackwell, 1987.

MALINVAUD, E. Econometric methodology at the Cowles Commission: rise and maturity. In: KLEVORICK, A. K. (Ed.). **Abstracted from the Cowles fiftieth anniversary volume**. New Haven: Cowles Foundation for Research in Economics, 1983. Retrieved Mar. 26, 2024, from: https://cowles.yale.edu/sites/default/files/2022-08/50th-malinvaud.pdf.

MALINVAUD, E. L'économétrie dans l'élaboration théorique et l'étude des politiques. **L'Actualité Économique**, v. 73, n. 1-2-3, p. 11-25, 1997.

MALINVAUD, E. La modélisation en macroéconomie appliquée: quarante ans après. **Brussels Economic Review**, v. 160, p. 329-342, 1998.

NUNES, A. F. N. de; PORTUGAL, M. S. Intermediários financeiros e política de crédito em um modelo DSGE para o Brasil. **Revista Brasileira de Economia**, v. 72, n. 3, p. 361-390, 2018.

PHILLIPS, A. W. The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861-1957. **Economica**, v. 25, n. 100, p. 283-299, 1958.

REISS, J. **Error in economics**: towards a more evidence-based methodology. London: Routledge, 2008.

ROMER, P. The trouble with macroeconomics. **Journal Title**, p. 1-20, 2016. Retrieved Mar. 10, 2021, from: https://ccl.yale.edu/sites/default/files/files/The%20Trouble%20with%20Macroeconomics.pdf.

SARGENT, T. J. The observational equivalence of natural and unnatural rate theories of macroeconomics. **Journal of Political Economy**, v. 84, n. 3, p. 631- 640, 1976.

SARGENT, T. J. **Dynamic macroeconomic theory**. Cambridge, United States: Harvard University Press, 2009.

SARGENT, T. J.; WALLACE, N. Rational expectations and the dynamics of hyperinflation. **International Economic Review**, v. 14, n. 2, p. 328-350, 1973.

SERGI, F. **DSGE models and the Lucas critique**: a historical appraisal. Bristol: UWE Bristol, 2018. (Economics Working Paper Series, n. 1806).

SIMS, C. A. Money, income, and causality. **The American Economic Review**, v. 62, n. 4, p. 540-552, 1972.

SIMS, C. A. Macroeconomics and reality. **Econometrica**, v. 48, n. 1, p. 1-48, 1980a.

SIMS, C. A. Comparison of interwar and postwar business cycles: monetarism reconsidered. **The American Economic Review**, v. 70, n. 2, p. 250-257, 1980b.

SIMS, C. A.; GOLDFELD, S. M.; SACHS, J. D. Policy analysis with econometric models. **Brookings Papers on Economic Activity**, v. 1982, n. 1, p. 107-164, 1982.

SMETS, F.; WOUTERS, R. Shocks and frictions in US business cycles: a Bayesian DSGE approach. **The American Economic Review**, v. 97, n. 3, p. 586-606, 2007.

SOLOW, R. M. **Growth theory**: an exposition. Oxford: Clarendon Press, 1970.

SUMMERS, L. H. The scientific illusion in empirical macroeconomics. **The Scandinavian Journal of Economics**, v. 93, n. 2, p. 129-148, 1991.

THEIL, H. **Economic forecasts and policy**. Amsterdam: North-Holland, 1958.

TINBERGEN, J. **An econometric approach to business cycle problems**. Paris: Hermann and Cie, 1937.

TOBIN, J. **Asset accumulation and economic activity**: reflections on contemporary macroeconomic theory. Chicago: University of Chicago Press, 1982.

VEREDA, L.; CAVALCANTI, M. A. F. H. **Modelo dinâmico estocástico de equilíbrio geral (DSGE) para a economia brasileira**: versão 1. Brasilia: Ipea, Mar. 2010. (Texto para Discussão, n. 1479).

COMPLEMENTARY BIBLIOGRAPHY

LINDÉ, J. DSGE models: still useful in policy analysis? **Oxford Review of Economic Policy**, v. 34, n. 1-2, p. 269-286, 2018.