

**Friedman, Becker, and Klein on Statistical Illusions:
Devising Criteria to Judge the Performance of
Large-Scale Macroeconometric Models**

Erich Pinzón-Fuchs

Lecturas de Economía - No. 98. Medellín, enero-junio 2023



Erich Pinzón-Fuchs

Friedman, Becker, and Klein on statistical illusions: Devising criteria to judge the performance of large-scale macroeconomic models

Abstract: *In 1957, Milton Friedman and Gary S. Becker accused Keynesian modelers of falling into a statistical illusion based on their inappropriate treatment of the consumption function and their inability to yield accurate predictions of income. However, despite their criticism, Friedman and Becker (1957), albeit unconsciously, significantly contributed to Keynesian large-scale macroeconomic modelling, anticipating both the dissemination of a method to evaluate model performance and the important role that dynamics would play in the construction of these models. This method was full model simulation, later routinized as computer dynamic simulation. Here, I provide an account of the early discussions on the evaluation of the performance of macroeconomic models and how ideas about dynamics entered the discussion of macroeconomic modelling.*

Keywords: *History of macroeconomics, history of econometrics, computer dynamic simulation, large-scale macroeconomic modelling, Lawrence R. Klein, Milton Friedman.*

JEL Classification: B22, B23, B40.

Friedman, Becker y Klein y las ilusiones estadísticas: desarrollando criterios para juzgar el desempeño de los modelos macro-económicos de gran escala

Resumen: *En febrero de 1957, Milton Friedman y Gary Becker (1957) acusaron a los modelizadores keynesianos de caer en una ilusión estadística basada en la inadecuada utilización de la función de consumo en sus modelos y su inhabilidad de producir predicciones acertadas del nivel futuro de ingreso. Sin embargo, a pesar de sus críticas, Friedman y Becker (1957) inconscientemente terminaron contribuyendo de forma importante a la modelización macroeconómica keynesiana, anticipando, por un lado, la diseminación de un método para la evaluación del desempeño de los modelos keynesianos y, por otro lado, el rol que la dinámica jugaría en la construcción de este tipo de modelos. El método anticipado fue la simulación completa de modelos, más tarde convertida en la rutinaria simulación dinámica computarizada. Este artículo cuenta la historia de las discusiones tempranas sobre la evaluación del desempeño de los modelos macroeconómicos keynesianos y la forma en que las ideas sobre la dinámica entraron en la discusión de la modelización macroeconómica.*

Palabras clave: *historia de la macroeconomía, historia de la econometría, simulación dinámica computacional, modelización macroeconómica de gran escala, Lawrence R. Klein, Milton Friedman.*

<https://doi.org/10.17533/udea.le.n98a349111>



Este artículo y sus anexos se distribuyen por la revista *Lecturas de Economía* bajo los términos de la Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0. <https://creativecommons.org/licenses/by-nc-sa/4.0/>

Friedman, Becker et Klein et les illusions statistiques: développer des critères pour juger la performance des modèles macroéconomiques à grande échelle

Résumé: *En février 1957, Milton Friedman et Gary Becker (1957) accusaient les keynésiens de tomber dans une illusion statistique basée sur l'utilisation tout à fait inappropriée de la fonction de consommation dans leurs modèles, ainsi que leur incapacité à produire des prédictions précises sur l'évolution du niveau de revenu futur. Malgré leurs critiques, Friedman et Becker (1957) ont fini, involontairement, par apporter une contribution importante à la modélisation macro-économétrique keynésienne. D'une part, ils ont participé à la diffusion d'une méthode d'évaluation des performances des modèles keynésiens et, d'autre part, ils ont mis en évidence le rôle que jouerait la dynamique dans la construction de ce type de modèles. La méthode était la simulation complète du modèle, convertie plus tard en simulation dynamique informatique. Cet article raconte l'histoire des premières discussions sur l'évaluation de la performance des modèles macro-économétriques keynésiens et comment les idées sur la dynamique ont fini pour rentrer dans la discussion sur la modélisation macro-économétrique.*

Mots clés: *histoire de la macroéconomie, histoire de l'économétrie, simulation computationnelle dynamique, modélisation macro-économétrique à grande échelle, Lawrence R. Klein, Milton Friedman.*

Cómo citar / How to cite this item:

Pinzón-Fuch, E. (2023). Friedman, Becker, and Klein on statistical illusions: Devising criteria to judge the performance of large-scale macroeconomic models.. *Lecturas de Economía*, 98, 131-165. <https://doi.org/10.17533/udea.le.n98a349111>

Friedman, Becker, and Klein on statistical illusions: Devising criteria to judge the performance of large-scale macroeconometric models

Erich Pinzón-Fuchs ^a

–Introduction. I. –The Controversy’s milieu: Marshallian and Walrasian approaches to macroeconomic modelling. –II. Economics as an empirical and modelling science: a common background. –III. The “simple Keynesian model” and the Controversy on statistical illusions. –IV. Keynesian Responses to Friedman and Becker. –V. Brown’s Anticipation of the Permanent Income Hypothesis. –VI. Differences in the Concepts of Prediction. –VII. Klein’s Emphasis on Extrapolation. –VIII. Full Model Simulations as a Turing Test. –Conclusions. –Acknowledgements. –References.

Original manuscript received on 12 March 2022; final version accepted on 26 November 2022

Introduction

In the 1957 February issue of the *Journal of Political Economy*, Milton Friedman and Gary S. Becker published their paper “A Statistical Illusion in judging Keynesian Models”. In this work, their purpose was to question Keynesian macroeconometric models for their inappropriate treatment of the consumption function, and for their inability to yield accurate predictions of income. Friedman and Becker (1957, p. 64) claimed that in replacing “the ultimate objective of predicting income [...] by the proximate objective of predicting consumption”, Keynesian modelers had fallen into a statistical illusion. According to them, the illusion resulted from the adoption of the relative error in predicting consumption as a criterion to judge the performance of macroeconometric models. In a nutshell, Friedman and Becker’s (1957) criticism of the “statistical illusion” in judging Keynesian models consisted of two points. While their first point focused on how to correctly evaluate model performance, their second point was related to the specification of the consumption function itself.

^a *Erich Pinzón-Fuchs*: Professor at Universidad Nacional de Colombia, Bogotá, Colombia.
E-mail: erapinzonfu@unal.edu.co. <https://orcid.org/0000-0002-3912-9907>

John Johnston (1958), Edwin Kuh (1958), and Lawrence R. Klein (1958) provided three separate responses in defence of the Keynesian approach, which comprised four common claims. First, the Keynesian consumption function used in the critique was at best a pedagogical device and did not make justice to the functions used in actual modelling. Second, Friedman and Becker had discovered nothing new and were raising problems already known to econometricians. Third, Friedman and Becker's results could be easily improved through the use of alternative consumption functions. Fourth, there had been no statistical illusions in judging Keynesian models.

The controversy on statistical illusions was not an isolated event in the history of macroeconometrics, though. This controversy was embedded within a larger debate that echoed discussions from previous decades and anticipated other discussions to come. This larger debate opposed two research programs, each of which claimed to provide the best empirical approach to economics (see Boumans, 2016; Pinzón-Fuchs, 2016; 2017). More specifically, this debate took place between the statistical economics approach stemming from the tradition of the National Bureau of Economic Research (NBER), and the econometrics program inspired, among others, by the works of the Cowles Commission.

As is often the case for scientific debates, the nature of the issue underlying the controversy on statistical illusions was methodological (although technical and theoretical matters were also discussed). On this occasion, important questions were raised about the establishment of a rigorous criterion to evaluate the performance of macroeconomic models, which would improve both model specification and predictions. Indeed, a powerful criterion to evaluate model performance would guide the macroeconomist in her choice and adjustment of the model structure. Apart from its function as model selector, this criterion would also act as an instrument to shape and “mould” (Boumans 2005, pp. 13) macroeconomic models or theory.

Surprisingly enough, Friedman and Becker (1957) made an important contribution to large-scale macroeconomic modelling, anticipating both the dissemination of a method to evaluate model performance and the important role that dynamics would play in the construction of these models.

This method was full model simulation, later routinized as computer dynamic simulation.¹ Through their insistence on the use of the *relative error in the prediction of aggregate income* as the adequate criterion to evaluate model performance (instead of the relative error in prediction of aggregate consumption), Friedman and Becker advocated for the evaluation of structural equations to rely on *full model simulations* and not on predictions based on single equations. Accordingly, in order to evaluate their models, macroeconometricians should emphasize the simultaneous equation characteristics of structural equations (in particular on the accuracy of their predictions) rather than on their single equation characteristics (Bodkin 1995, pp. 53-54).

Yet, by the 1950s, macroeconometricians were not necessarily convinced about the requirement of this criterion nor did they have the computational power to bring about the burdensome calculations that would allow them to meet this criterion. Some considered, instead, that *out-of-sample predictions* or *extrapolations* (even when performed on single equation bases) would be more adequate to evaluate models than *sample predictions* (even if performed as full model simulations). While macroeconometricians' idea of prediction focused on what Klein (1958, p. 543) called the "painful [test] of experience" (i.e., the comparison between observed values (*ex-ante* or *ex-post*) and out-of-sample "concrete forecasting results" of alternative models), full model simulation could be understood as a sort of Turing test, or as an "imitation game" (Turing, 1950).

The increasing availability and improvement of computational methods, as well as the expansion of large-scale macroeconomic modelling in general, provided some of the necessary conditions for the dissemination of full model simulations (Ando & Modigliani, 1969; Bodkin, 1995; Bodkin et al., 1991). The successful application of this kind of simulations was also an important push in this direction. The particular case of Irma and Frank Adelman's

¹ I say "surprisingly enough" primarily because of Friedman and Meiselman's (1963) later presentation of their simple reduced form equation to compare the effects of fiscal and monetary policy that prolonged and heated up the debate between Keynesians and Monetarists. However, it also refers to Friedman's earlier opposition vis-à-vis the econometrics program, dating back to the 1940 s. On this last point, see Boumans (2016) and Pinzón-Fuchs (2016; 2017).

(1959) simulation of the Klein-Goldberger (1955) model certainly exerted an important effect on the establishment of this method as a standard criterion to evaluate model performance.

Although simultaneous equations characteristics represent only some among a whole battery of criteria used to judge models' performance, this controversy illustrates the way macroeconometricians discussed, developed, and adopted standards to judge (and improve) the performance of their models. My main claim is that independently of Friedman and Becker's critical claims and clear methodological differences regarding the Keynesian modelers, a common background of general principles turned this controversy (and the larger debate) into a fruitful and revealing episode of the history of macroeconomics. This common background consisted of the belief of three points by contemporary macroeconomists: (1) that they should adopt an empirical approach to economics, (2) that macroeconomic modelling was the way to go if they wanted to integrate rigorous thinking to economics, and (3) that statistical techniques were considered necessary to provide rigorous criteria to judge model performance.²

In this paper, I give an account of the early discussions on the evaluation of the performance of macroeconomic models and on the way ideas about dynamics entered the discussion of this, by the time, novel scientific practice. I begin by revisiting a fundamental difference between two modelling strategies in macroeconomics: the Marshallian approach adopted by Friedman and the Walrasian approach adopted by Klein. Then, I describe the common background shared by contemporary macroeconomists who saw their discipline already in the late 1950s necessarily as an empirical discipline that should be based on modelling and on the use of the latest statistical techniques. In the next two sections I describe in detail the Controversy on Statistical Illusions, I revisit the three Keynesian responses to Friedman and Becker's critique, before going into the details of one of the Keynesians' counterattacks which consisted on Klein's claim that Friedman's Permanent Income Hypothesis had been anticipated by T.M. Brown in the early 1950s. The last two sections

² I approach this controversy inspired by the "Strong Programme in the Sociology of Knowledge" following its four tenets as explained by David Bloor (1976, p. 7): causality, impartiality, symmetry, and reflexivity.

of the paper, I emphasize on Klein's response that the most important feature of his macroeconomic models consisted in their capacity to perform policy evaluation, and on their ability to mimic the economy, its movements, and predict turning points. Given the important role that macroeconomic modelling played in the spread of the use of the computer in economics, I suggest that the scale to which full model simulation of macroeconomic models was compared was an idea similar to the Turing Test, in which the model should be able to imitate reality. I present the conclusions in the last section of the paper.

I. The Controversy's milieu: Marshallian and Walrasian approaches to macroeconomic modelling

The controversy on statistical illusions was embedded within a larger methodological debate confronting two empirical approaches to economics: the "statistical economics" and the "econometrics program". While the NBER appeared as the most visible and important stronghold of the statistical economics approach, the Cowles Commission, during its years in Chicago (1939-1955), constituted the bastion of the econometrics program.³ Since at least the mid-1940s, Friedman had sustained longstanding discussions with various members of Cowles in which he had expressed his concern about the necessity of finding a sound and rigorous criterion to evaluate model performance.⁴ In fact, this was one of the main arguments in his 1946 debate with Oskar Lange (Friedman, 1946).

³ During the 1950s, however, Friedman (in the case of the NBER) and Klein (in the case of the econometrics program) established themselves as the two leading figures of these approaches, embodying a reinterpretation of the ancient Walras-Marshall divide as two alternative ways to build macroeconomic models.

⁴ It is worth noting that Friedman took the role of the critic in most cases. Klein (and the Cowles's members), instead, acted in general as defenders. This might have been the case, because large-scale macroeconomics enjoyed a "dominant" position only from the 1950s onwards. During the 1940s, however, the econometrics program was on a preliminary state, and the NBER enjoyed a particularly good position within the economics community. Judging from the *Measurement without theory* controversy between Tjalling C. Koopmans and Rutledge Vining, critics and defenders would have switched positions at the end of the 1940s.

Departing from a situation of involuntary unemployment, Lange's *Price Flexibility and Employment* (1944) investigated whether a decrease in money-wages could re-establish full employment using, according to Friedman (1946, p. 613), an "unreal and artificial" approach. Friedman's claim was not so much against abstraction, but rather about the way Lange used abstraction. According to Friedman, Lange attributed too much importance to the conformity of the model structure with the canons of logic, ignoring the model's "empirical application or test". Lange's use of "casual observation" (p. 618) as his only method to evaluate his model was not a sufficiently rigorous criterion to evaluate the relevance of the proposed functions, and it resulted in systems with no "solid basis on observed facts", yielding "few if any conclusions susceptible of empirical contradiction" (p. 619). Friedman's concern, again, was about Lange's (and the Commission's) lack of a rigorous method to evaluate model performance based on a sound empirical approach.⁵

Some important methodological works published during this period resulted, partly, because of the exchanges between Friedman and the Cowles's members. Friedman's criticisms ultimately led, even if unwittingly, to the strengthening of the Cowles's effort to provide a solid empirical approach to macroeconomic modeling.⁶ A recurrent subject in these works, notably in the case of Friedman's, related to the ancient Walras-Marshall divide. To Friedman, the important element of this divide did not consist of the classical opposition of general and partial equilibrium attributed to Léon Walras and Alfred Marshall. The ulterior motive of this divide relied on a more profound methodological question. Both Walras and Marshall understood that the economic system was fundamentally complex and that all its parts were interdependent. Thus, given this interdependency, the question was whether

⁵ Another important controversy between Friedman and the Commission took place at the occasion of the *NBER Conference on Business Cycles* in November 1949. This time, Friedman advocated, again, in favor of the establishment of a criterion to evaluate model performance. In particular, he proposed his "naive models" as standards (or null hypotheses) to compare the predicting performance of macroeconomic models. I will come back to this subject in the third section of this paper. For a more comprehensive account of this controversy, see Pinzón-Fuchs (2016, pp. 10-16).

⁶ See, for instance, Friedman (1946; 1949; 1953; 1955; 1958), Haavelmo (1943; 1944; 1947), Koopmans (1950; 1957), and Hood and Koopmans (1953) among other works.

the economist, interested in finding concrete solutions relative to some parts of the economy, should take the *entire system* into consideration (as Walras suggested) or focus only on a few parts of it (as Marshall proposed).

Friedman's response to this question pointed clearly in Marshall's direction. His Marshallian approach was based on the idea that economic theories – or models – should be perceived as a way to construct systems of thought or “engine[s] for the discovery of concrete truth” (Marshall, 1885, p. 159). Since both the economist's knowledge and his capacity to observe the world were partial, these systems should be constructed through the rigorous observation of specific parts of the economy and of the “real” world and not through elegant but empty abstractions of the entire system. In brief, these systems should provide “generalizations of the real world” instead of “formal models of imaginary worlds” (Friedman, 1946, p. 618). This approach, however, recognized that the economy was a complex system with interdependent relations, which had nothing to do with the notion of partial equilibrium.

Klein's Walrasian approach, on the contrary, was based on the idea that the economy should be considered as a whole despite the economist's inability to observe or understand the system in all its complexity. Independently of the economist's capacity to build a simple (or complex) model, his point of departure to build macroeconomic models should be the entire economic system. Klein's claim was that the combination of a mathematized framework of general equilibrium with statistical theory and with rigorous empirical work, would be useful to provide a tool of reasoning for understanding and intervening in the economy.⁷

It is worth noting that neither Friedman nor Klein explicitly alluded to the Walras- Marshall opposition during the controversy on statistical illusions. The larger methodological debate, however, was the background and the pillar of the controversy on statistical illusions and has to be born in mind at all times in the study of this controversy.

To be clear, I believe that there are two different levels in this controversy.

⁷ Through his illustration of what he calls “Cournot's Problem”, Kevin D. Hoover (1988, section 9.2) provides a clear exposition of this methodological divide between Marshall and Walras. See also Pinzón-Fuchs (2016, section II).

The first level, again, goes up to the larger debate (and to the opposition) between two alternative empirical approaches in macroeconomics: the “statistical economics approach” and the “econometric program”. At this level, given the nature of Friedman and Becker’s paper, there is no doubt that the authors insisted on the abandonment of the “econometric program”, even if they did so in a somewhat subtle way.⁸ It is at this first level, too, that Friedman would pick up his later debates with the Keynesian approach, in particular in Friedman and Meiselman (1963), and in Friedman and Schwartz (1963).⁹ At the second level, it seems as if Friedman and Becker had accepted to play within the rules of the “macroeconometricians’ game”, which was on its way to becoming the dominant approach in macroeconomics (see Bodkin et al. 1991). At this level, their critique could be qualified as constructive, for it sought to establish a criterion to evaluate model performance and improve model specification.

II. Economics as an empirical and modelling science: A common background

In his review of the controversy on statistical illusions, Ronald G. Bodkin (1995, p. 45) contended that any commentator on this discussion should “take into account [...] whether there was [...] a genuine disagreement, or whether the issues among the participants were principally semantic”.¹⁰ Perhaps this

⁸ “Subtle” compared to other papers, where Friedman was much more radical. See, for instance, Friedman’s (1951, 112) conclusion: “My own hunch [...] is that [...] attempts to proceed now to the construction of additional models along the same general lines [of the Cowles Commission] will, in due time, be judged failures” (see also Friedman 1946, 1953).

⁹ Given the scope of this paper I will not scrutinize further in this direction, but these later Keynesian- Monetarist debates should also be kept in mind in this discussion.

¹⁰ Judging from Friedman and Becker’s (1958a, 1958b) responses to Klein, Kuh, and Johnston alone, one possible interpretation of this controversy would indeed be that the differences between the participants were semantic and not substantial. In fact, Friedman and Becker considered the Keynesian responses “less [as] a criticism [...] than a valuable supplement” of their comment (1958a, p. 545). To them, “if [Klein, Kuh, and Johnston’s] comments nonetheless give the impression of being a criticism, it is [...] because, they interpret our note as having a different aim and different content than we intended it to have, and still think it has” (1958b, 298). The Keynesian modelers, however considered their responses neither

contention should be reformulated in a somewhat different way. There is no doubt that there were issues of substantial disagreement between the participants, as Bodkin recognized at the end of his paper.¹¹ More importantly, there was also a certain degree of agreement on fundamental points that allowed the participants to actually engage in a productive conversation. I claim that, in order to make a contribution to the history of macroeconometrics, anyone commenting on this controversy should take these points of agreement seriously into account, for these points allowed the participants to engage in a discussion in the first place.

It is clear that for any scientific controversy to take place and for it to be fruitful, some agreement must exist between its participants. This agreement, reflected in some general, but fundamental points, can be understood as a sort of “common background”. Without this common background, fertile scientific controversies would simply become fruitless and meaningless attacks or semantic misunderstandings at best. Thus, a brief characterization of this background is useful for at least two reasons: First, it sheds light on some of the important issues at stake, making them more visible; second, it provides the possibility of historicizing the controversy, putting some flesh on the methodological skeleton of the discussion.

In the particular case of the controversy on statistical illusions, no less than three fundamental points of agreement built the common background between Friedman and Klein (and the other participants). Both factions considered that:

1. Economics was a science whose foundations should be grounded on an empirical approach.
2. A way of integrating a rigorous empirical approach in economics was through the practice of econometric modelling (in a broad sense).¹²

as “supplements” of Friedman and Becker’s paper, nor as semantic misunderstandings. To them, there were substantial points to be defended.

¹¹ Bodkin (1995, p. 53) termed “the debate” these issues of substantial disagreement.

¹² Econometric modeling in a “large sense” means that the definition of econometrics is not reduced to the structural econometric approach developed at the Cowles Commission, but it can also embrace other definitions of econometrics like the statistical approach of the

3. Statistical techniques could help economists to develop rigorous criteria to judge the performance of their models, helping them in the discovery and further specification of the underlying economic mechanisms.

These general principles established the point of departure of the controversy between Friedman and Becker, and Klein. There are, of course, important (and sometimes irreconcilable) differences between the methodological positions of these authors concerning the way to tackle each of these principles. Yet, the important point here is that this common background was not exclusive to Friedman, Becker, and Klein, but was part of the economists' scientific community, and that discussions, criticisms, and controversies based on this common background contributed, even if unwittingly, to the devising and determination of empirical criteria to judge the performance of large-scale macroeconomic models.

III. The “Simple Keynesian Model” and the Controversy on statistical illusions

To illustrate their criticism, Friedman and Becker (1957, p. 64) used a Keynesian model “in its simplest form”, specified according to the absolute income hypothesis.¹³ This simple model contained a consumption function (1) and a national income accounting identity (2). Investment was considered autonomous:

$$C_t = a + bY_t + \dots + u_t \quad (1)$$

$$Y_t = C_t + I_t \quad (2)$$

with C_t and Y_t the aggregate consumption and income in a given year, while a and b are parameters, u_t is the stochastic perturbation of the consumption function, and I_t is realized investment, which, by definition, is also realized savings.

NBER. See Boumans and Dupont-Kieffer (2011) for a further discussion of the definition of econometrics in a larger sense.

¹³ According to John Maynard Keynes's (1936/1965, p. 96) formulation, this means that “men are disposed, as a rule and on the average, to increase their consumption as their income increases, but not by as much as the increase in their income”.

The common forecasting practice consisted of the use of the reduced forms of the model. In this case, the reduced form equations (3) and (4) are obtained by replacing the structural equation (1) in the accounting identity (2):

$$Y_t = \frac{a}{1-b} + \frac{1}{1-b}I_t + \frac{u_t}{1-b} \quad (3)$$

$$C_t = \frac{a}{1-b} + \frac{b}{1-b}I_t + \frac{u_t}{1-b} \quad (4)$$

Forecasts of the level of consumption obtained through the reduced form equation (4), would present a small relative error of prediction V_c (see the first column of Table 1), according to which macroeconometricians (under the statistical illusion) tended to consider that their model, as a whole, would perform well.¹⁴ This, however, was a misleading consideration, since conclusions on the performance of the whole model were drawn based on single equation characteristics. This procedure, of course, was not necessarily consistent and entailed misleading results. For instance, Friedman and Becker (1957) showed that, if macroeconometricians kept this (incorrectly) specified model structure, forecasts of the level of income would present dramatically higher relative errors of prediction, revealing that the model, as a whole, did not perform well. This was a relevant point, since the ultimate objective of macroeconometricians was, indeed, to forecast the level of income, while forecasting the level of consumption was only a secondary objective.

Based on the data published in Raymond Goldsmith's *A Study of Savings* (1955), Friedman and Becker performed a numerical exercise to illustrate their point and to compare the performance of six different consumption equations following four distinct procedures of prediction.¹⁵ Table 1 describes the relative error in prediction of these consumption functions.

¹⁴ To measure the relative error in predicting C , the standard deviation of the errors σ_u is divided by \bar{C} (the average value of C): $V_c = \frac{\sigma_u}{\bar{C}} = \frac{\sigma_u}{1+b+\bar{Y}}$, with \bar{Y} the average value of Y . The quantity V_c can be understood as a measurement of the "usefulness" of equation (1) as a predictor of consumption.

¹⁵ The alternative consumption functions are presented in annex 1 and are given here under equations (7)-(12). They are equivalent to Friedman and Becker's (1957) equations (12)-(17),

Table 1. *Errors in the Prediction of Consumption and Income from Different Consumption Functions.*

CONSUMPTION FUNCTION	RELATIVE ERROR IN PREDICTING			
	Consumption from a Known Level of Income	Income from a Known Level of Investment	Income from an Estimated Level of Investment* (Consumption Known with Certainty)	Income from Estimated Consumption and Investment*
	(1)	(2)	(3)	(4)
(12) $C(t) = 70.2 + .780Y(t) + u_1$049	.196†	.100	.242‡
(13) $C(t) = .890Y(t) + u_2$057	.462†	.200	.551‡
(14) $C(t) = (.879)(.4) \int_{-\infty}^t e^{-.38(T-t)} Y(T) dT + u_3$040	.051	.029	.059
(15) $C(t) = 162.0 + 237.0e^{-.02t} + u_4$066	.059	.020	.062
(16) $C(t) = C(t-1) + u_5$068	.061	.020	.065
(17) $C(t) = e^{.02}C(t-1) + u_6$064	.057	.020	.061

* The relative error in predicting investment is assumed to be .20.

† Instead of .196 and .462, these figures would be .147 and .420, respectively, if more efficient estimates of the parameters of these consumption functions had been used.

‡ Instead of .242 and .551, these figures would be .178 and .465, respectively, if more efficient estimates of the parameters of these consumption functions had been used.

Source: Friedman and Becker (1957, p. 67).

In predicting consumption from income (column 1) the Keynesian functions, equations (12) and (13), presented a small relative error of prediction (0.49 and 0.57, respectively), showing superior results (in terms of smaller relative errors) compared to all the other consumption functions, except for (14), whose relative error is the smallest of all (0.40). This was the source of the statistical illusion that would misleadingly justify the maintenance of the Keynesian consumption function.

As opposed to the conclusions that could derive from the evidence that was first visible (the small relative error in predicting consumption), the simple

reproduced here as Table 1. The six alternative consumption functions can be grouped in three categories: (7)-(8) are “Keynesian functions”, (9) is a function specified according to the permanent income hypothesis, and (10)-(12) are naive models. Also, note that the consumption functions (8) and (9) do not take into account the autonomous part of consumption.

Keynesian model actually presented important flaws in its consumption function. These flaws were translated into misspecifications that led the model to perform poorly in its forecasts of the level of income (see column 2 of Table 1) and was the result of the interaction of single equation random errors. “[A] given error in predicting consumption [was] magnified by the multiplier process into a much larger error in predicting income” (Friedman and Becker 1957, p. 66). This was so, because “the accuracy of the estimate of income depends not only on the accuracy of the estimate of consumption *but also on the form of the consumption function* – in particular, on the fraction of consumption which [the consumption function] designates as ‘autonomous’” (p. 64, my emphasis). The problem was that the statistical illusion would hide the misspecification of the whole model and that small relative errors obtained from the single equations built up to produce even larger relative errors in predicting other variables of the model. The presence of multipliers in the Keynesian models made these errors even more important.

This argument set the basis for Friedman and Becker (1957, p. 75) to conclude that “improvement in the consumption function [...] may [...] make a greater contribution to our ability to predict income than improvement in the estimates of investment”. The problem of Keynesian models was not one of statistical or mathematical adequacy, but one of “the structure [these models] attribute to the economic system” (p. 73). Since Keynesian modelers did not apply the correct criterion for judging model performance, they fell into a statistical illusion that led them to accept a misleading model specification, hence the wrong structure. Friedman and Becker did not limit their claim to this already important point but went further to claim the general principles to be taken into account to specify a correct consumption function. These principles, which were, of course, related to Friedman’s (1957) *A Theory of the Consumption Function*, his Marshallian approach, and his “rule of parsimony”, (Klein, 1958, p. 545) were put forward in order to build macroeconometric models.

It is worth noting, that even if the particularly simple Keynesian system used in the critique constituted a pedagogical tool only, the predicting procedure denounced by Friedman and Becker was a common practice among Keynesian macroeconometricians during the 1950s (Bodkin 1995, p. 46).

Friedman and Becker were then correct both in criticizing this practice and in raising this important methodological concern. Indeed, the fact that the relative errors in predicting the level of consumption were small, was not a sufficient condition to consider the whole model accurate in predicting income. Statistical adequacy based on the prediction of a single equation did not immediately mean that the complete model would perform well.

IV. Keynesian Responses to Friedman and Becker

Johnston (1958), Kuh (1958), and Klein (1958) provided three separate responses to Friedman and Becker (1957). Although each one referred to different important aspects of the controversy, four common claims were made. The first reflected a shared feeling that the macroeconometric approach had been unfairly criticized. Indeed, given the obsolete nature of the functions and the time period (1905-1951), to which the simple model had been fitted in the critique, Johnston (1958) declared that it was “hardly [...] surprising to read the final conclusion that these simple functions give an inadequate specification of dynamic characteristics of the economic structure” (pp. 296-297). According to Johnston (Kuh, and Klein), and this was the second common argument, the type of functions used in the critique had been abandoned in actual practice at least since the last ten years, and in particular after the publication of Franco Modigliani’s (1949) paper, “Fluctuations in the Saving-Income Ratio: A Problem in Economic Forecasting”.

In this paper, Modigliani criticized Jacob Mosak (1945) and Arthur Smithies’s (1945) forecasting methods because they relied on simple Keynesian functions similar to the one used by Friedman and Becker (1957) to illustrate the Keynesian approach. Modigliani highlighted the difficulties related to the use of these “simple” functions in economic forecasting. Indeed, economists faced a complex task in their forecasting objective, given the “pronounced discrepancy between the cyclical, or short-run, and the secular, or long-run” forms of economic relations in a period of “violent cyclical fluctuations” like the first half of the twentieth century (particularly in the US). Notably, Modigliani suggested a method “by which it seem[ed] possible to estimate” both the secular and the cyclical forms of economic relations (p. 371). Based

on Modigliani (1949), Johnston (1958) estimated two consumption functions that allowed for the differentiation of (a) the short-run or cyclical marginal propensity to save, and (b) the long-run average and marginal propensity.

$$C_t = a + bY_t + Y_t^0 \quad (5)$$

$$C_t = a + bY_{t+1} + Y_t^0 \quad (6)$$

where Y_t^0 is the highest income experienced before time t .

Judged according to Friedman and Becker's (1957) criteria, both equations (5) and (6) presented a well predicting performance. The relative error in predicting consumption and income from equation (5) was $V_C = 0.0304$, and $V_Y = 0.053$, respectively. In the case of equation (6), these relative errors were of $V_C = 0.53$, and $V_Y = 0.0485$. With this estimation, Johnston not only made his point that the simple model used by Friedman and Becker as an illustration of the Keynesian model was no longer used in actual practice but also that alternative consumption functions that could easily improve the results obtained by Friedman and Becker existed (this was the third common claim of the three respondents). This improvement of the consumption functions did not consist of envisaging a future possibility of refining existing functions. In fact, far more sophisticated consumption functions which yielded more adequate results, even according to Friedman and Becker's criteria, had already been in use. In this respect, Klein (1958, p. 540) asserted that in the actual practice, those "who have seriously attempted forecasting from Keynesian models had used much more complicated systems in which the reduced-form equation for consumption is vastly different from and [...] superior to [Friedman and Becker's]".

The fourth common claim consisted of the macroeconometricians' consideration that there had been no statistical illusion. The three respondents did recognize, however, the existence of simultaneous equations biases in the forecasts obtained from reduced form equations. Yet, these problems were already known to econometricians and, according to the respondents, Friedman and Becker had discovered nothing new. Klein, for instance, claimed that "when Friedman and Becker showed that the variance of $[\frac{u_t}{1-b}]$ in the multiplier equation is much larger than the variance of $[u_t]$ in the consumption

equation (both expressed as a percentage of average consumption or income), [Friedman and Becker] are not proving the general inadequacy of Keynesian models or even of the consumption function". To Klein, they were "simply giving a laborious demonstration of the fact, already well known, that the simple multiplier model is not suitable for more than pedagogical use in the classroom" (p. 540) and acknowledging relative errors in prediction would accumulate in models presenting important multiplier effects.

Although this would concede the argument to Friedman and Becker if seen from a purely methodological perspective, in his response, Klein seemed to intentionally deviate attention from it by providing a different interpretation of what the critique actually meant. Referring to Haavelmo (1947, pp. 539-540), Klein claimed that "the appropriate way to estimate the structural parameters of equations in a complete system is to regard the whole set of equations simultaneously from a statistical point of view". In the case of the consumption function, this meant that "the propensity to consume should be estimated from the reduced forms and not by direct reference to the consumption function, to the neglect of the accounting definition". Since Klein's response seems to accept Friedman and Becker's criticism, one understands why Friedman and Becker (1958, p. 545) considered the econometricians' responses (and especially Klein's) as a "supplement" of their article. Still, the problem raised was that, although Klein was probably right in evocating the strength of his models, he focused too much on the ability of his reduced forms to yield accurate *extrapolations* of the economy. Klein, however, misses the point that this criterion does not always constitute a sufficient measure of the accuracy of model performance.

V. Brown's Anticipation of the Permanent Income Hypothesis

Another important point of the controversy has to do with Klein's (1958, p. 541) claim that "Brown's work on lags in consumer behavior is truly a complete anticipation of the Friedman-Becker article".¹⁶ T. M. Brown (1952)

¹⁶ Klein quoted Christ (1951), Pairs (1956), Koyck (1954), Klein and Goldberger (1955), and Stone and Rowe (1956) also as anticipating Friedman and Becker's proposed consumption function: $C(t) = k\beta \int_{-\infty}^t e^{(\beta-a)(T-t)} Y(T) dT + u_3$.

had developed his “habit persistence theory”, where, contrarily to Modigliani (1949), previous real consumption, and not previous income played an important role in the determination of the level of consumption. Brown (1952) claimed that:

[The] lag effect in consumer demand was produced by the consumption habits which people formed as a result of past consumption. The habits, customs, standards, and levels associated with real consumption previously enjoyed become ‘impressed’ on the human physiological and psychological systems and this produces an inertia or ‘hysteresis’ in consumer behaviour (p. 539).

A similar feature between Brown (1952) and Modigliani (1949) was that both approaches allowed for distinguishing between short-run and long-run marginal propensities to consume.

In his “supplementary comment”, Friedman (1958) actually conceded this point to Klein (1958, p. 549) by stating that he “did not recognize that [his] procedure was equivalent to Brown’s use of the consumption of the preceding year, and [that] Klein [was] quite right in criticizing [him] for this error of omission”. As noted by Bodkin (1995, p. 55), Friedman might have hastened too much in conceding this point of the controversy to Klein. Years later, Balvir Singh and Aman Ullah (1976, pp. 101-102) showed that “Brown’s model [was] by no means ‘truly a complete anticipation of Friedman’”. In fact, “even though the Friedman model [...] looks quite similar to the one suggested by Brown, the models differ in terms of the nature of the regressors, interpretation of the error term, and the nonlinearity in parameters”.

In my opinion, the essential point here is not whether Brown anticipated Friedman. The truly important point is whether the Keynesian and the non-Keynesian models present specifications close enough to be considered similar. This is important in order to understand the extent to which other criteria to judge model performance and model relevance, apart from the methodological approach itself, play a role in moulding and shaping economic models. Economists, like other scientists, rely on a battery of criteria to choose among different model specifications and to shape their models or theories.

This battery of criteria is not completely rigid and might be composed of a variety of elements such as statistical tests, adequacy of mathematical forms and procedures, the nature of data, political convictions, or even particular material and historical conditions like the existence and availability of digital computers, or the access to funding. Indeed, this battery of criteria belongs to the common background shared by macroeconomists at the time, whether Keynesians or not in their methodological approach. In this sense, these criteria to judge model performance are defined by the common background of the community, by their common practices and beliefs, rather than by their affiliation to a particular methodological or theoretical approach.

VI. Differences in the Concepts of Prediction

In terms of methodology, the most important point discussed in the controversy was about how to adequately evaluate model performance. Macroeconometricians considered that *ex-ante* or *ex-post* forecasts or *extrapolations* from the reduced form equations were an adequate way to evaluate model performance. As discussed above, both the critics and the econometricians knew that this procedure could produce biased results. Keynesians thought of these biases as the result of a statistical problem, particularly simultaneous equations biases. Friedman and Becker (1957), however, insisted that the biases were caused by a poor model specification, considering them a result of an economic theoretic problem. According to the Keynesians, the best way to avoid these problems was through painstaking tinkering, re-estimation, and actualization of the model, and through the inclusion of more explanatory variables in each equation. Klein (1958, 545) for instance, claimed that “researchers [had] sought improvement in the Keynesian consumption function through the introduction of new variables” and that there were “great limits to the extent to which one can come upon radically improved results by juggling about the same old variables in a different form”:

Instead of adhering to the “rule of parsimony”, we should accept as a sound principle of scientific inquiry the trite belief that consumer economics, like most branches of our subject, deals with complicated phenomena that are not likely to be given a simple

explanation [...] The addition of extra predetermined variables (not lagged incomes) that are not correlated with income or that do increase the multiplier are likely to improve the fit of the multiplier equation at the same time that they are improving the fit of the consumption equation. I venture to predict that much good work will be done in the years to come on adding new variables to the consumption function and that it will not be illusory.

Friedman and Becker (1957), on the contrary, made a plea for econometricians to adopt a different economic structure, based on a consumption function specified according to the permanent income hypothesis. Their plea, nonetheless, was also directed towards the development of parsimonious models that would explain more through simpler equations. In this particular case, both critics and respondents were exclusively focusing on how to solve the problem of obtaining biased results. However, none of them actually attempted to identify if the origin of the problem was statistical or economic.

To detect and define the nature of the problem, Keynesians and critics had to find an adequate measure, allowing for the understanding of the kind of problem they were facing. In this case as well, they still could not agree on what an adequate criterion would be since they valued different kinds of *predictions*. On the one hand, the Keynesians thought that *extrapolation* or *out-of-sample predictions* were a more important kind of *prediction* in that it constituted the more adequate criterion to judge model performance. Extrapolation meant that one should test the model's theoretical and statistical whole structure not only to fit the observed data but also to predict values that were outside the observed sample. Yet, by the 1950s, these extrapolations were performed only on the bases of reduced form equations and not on the bases of whole model simulations, partly because computers were still not available for economists in general, but more importantly, because macroeconometricians believed that reduced forms forecasts were adequate criteria to judge model performance.

On the other hand, Friedman and Becker (1957) suggested that predicting did not necessarily mean extrapolating. To them, retrodiction and extrapolation were equivalent in evaluating model performance, and the truly

important criterion was to find out whether extrapolation or retrodiction was performed on the basis of reduced form equations or on the basis of full model simulations. Once again, to them, only full model simulations would provide adequate ways of evaluating model performance.

VII. Klein's Emphasis on Extrapolation

A somewhat surprising way of interpreting Friedman and Becker's (1957) paper is to consider it as an important contribution to large-scale macroeconomic modelling. In fact, their critique can be understood as an anticipation of full model simulation as an important method to evaluate model performance. In this sense, again, the evaluation of the simultaneous equations characteristics of structural functions would be a better criterion to judge model performance compared to single equations characteristics.

Although Klein later adopted this criterion as the rest of the discipline did (Klein, 1969), in the 1950s, he was skeptical about Friedman and Becker's proposition. His skepticism relied on his different views on the nature of predictions. To Klein, Friedman and Becker's proposition put too much emphasis on the goodness-of-fit measures obtained for *sample period predictions* or *retrodictions*. Instead, Klein appraised other criteria like *out-of-sample predictions* or *extrapolations*, like the model's capacity to predict turning points in the economy. Making a clear reference to Friedman (1953), Klein considered that Friedman and Becker's (1957) proposition was rather "strange", since "Friedman [had], on many occasions, stressed the criterion of predictive ability as a suitable test of theory". And yet, in his defensive effort, Klein disregarded the fact that full model simulation provided also a kind of prediction. Only, this was a different kind of prediction.¹⁷

¹⁷ It is important to note also, that full (or even block) model simulations represented a huge technical challenge for econometricians, considerably increasing the bulk of their calculations. Furthermore, computers were not only in an early stage of development, but they were not available to every economist. As Irma Adelman (2007, p. 29) recalls, "the first application of digital computers to the solution of a macroeconomic model occurred through happenstance", since the computer actually belonged to the Berkeley radiation laboratory and since Irma's husband, Frank, was actually hoping to use the first general purpose computer (the IBM 650) in a physics problem. Yet, given "that all physics problems

Klein's idea of prediction focused on the "painful [test] of experience", i.e., the comparison between observed values and *ex-ante* or *ex-post* out-of-sample "concrete forecasting results" (p. 543) of alternative models. For instance, to Klein and Goldberger (1955, p. 72):

The severest test of any theory is that of its ability to predict. Our equation system presents a theory of economic behavior in the aggregate. We have fitted the model to the sample, and although it may be an achievement to find a structural system which does fit the observed facts, we cannot be satisfied with the performance of the system solely with reference to the sample data [...] In a broad sense, we mean, by prediction, the ability of the equations to explain aggregate economic behavior for sets of observations outside the sample.

Klein considered both *ex-ante* and *ex-post* extrapolations as important constituents of his predictions. These were complementary types of predictions, and both had to be conducted in order to evaluate model performance. Yet, between *ex-ante* extrapolations, Klein considered that policy simulation was more useful than mere forecasts of the level of particular endogenous variables. Klein and Goldberger, for instance, explained this procedure in the following way (1955, p. 72):

We have approached this problem empirically from two points of view. In an *ex-post* sense we may insert observed and essentially correct values of predetermined variables in the model and solve it algebraically for the values of endogenous variables in the forecast period. One interested in the degree to which our model represents a true picture of behavior should base this judgement of performance on this *ex-post* type of extrapolation. This is a case of testing the model outside the confines of the sample and determining how well it fits actual observations when there is no statistical forcing towards conformity [...] If the *ex-post* extrapolations have shown a system to give a good explanation of economic behavior, we can then place a measure of reliance on the use of this system to show what *would* happen

were too complex", they decided to look for a more "suitable model": the 25- equations Klein-Goldberger macroeconometric model.

if exogenous variables *were* to be placed at particular assumed levels. We might show, for example, the sensitivity of aggregate activity to variations in government tax-expenditure plans. This is a form of *ex-ante* forecasting and may be a more useful econometric application than pure forecasts of the levels of endogenous variables (emphasis in the original).

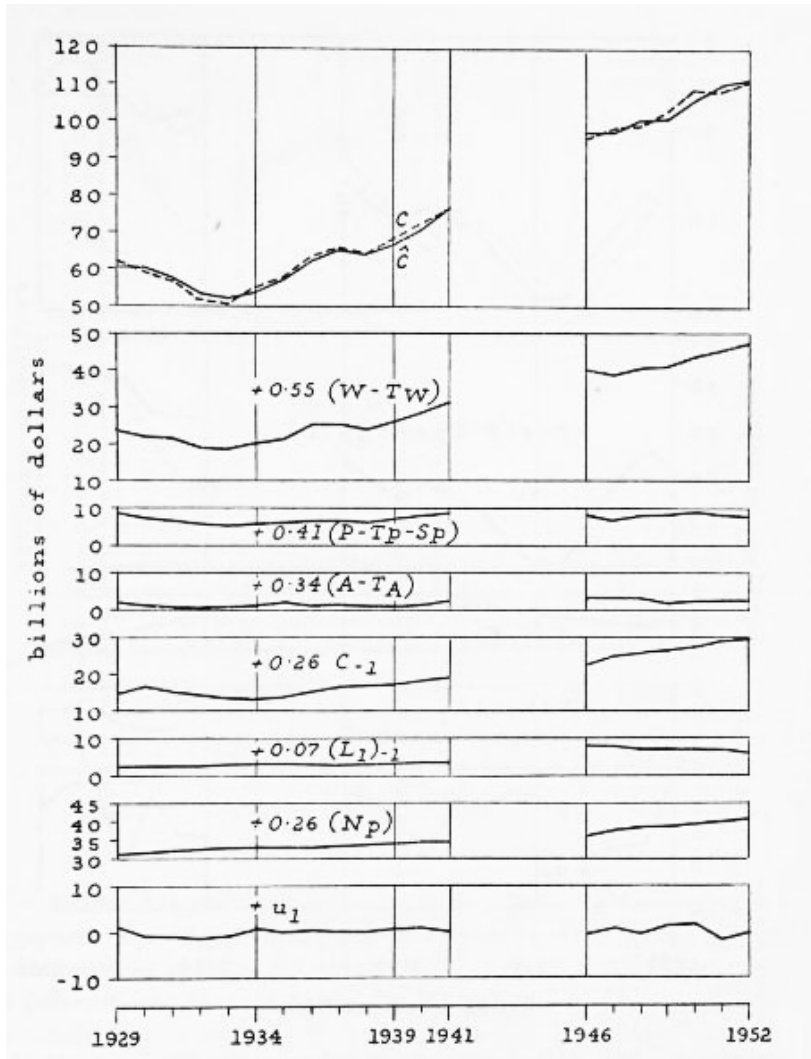
Figure 1 is an illustration of the type of ex-post forecasts performed by Klein and Goldberger (1955). Klein and Goldberger presented ten figures of this type, where they examined the behaviour of the ten most important equations of their model individually. Each graph consisted of a first chart (at the top of the figure) where the observed and forecasted trends of the endogenous variable for the whole period (1929-1952) were presented. The remaining lower charts presented the individual contributions of the exogenous variables and the errors in explaining the behaviour of the endogenous variable. Figure 1 reproduces the example of the consumption function. These variables were examined individually, in part, because of the computational burden that a full model simulation would mean for the time. However, with the advent of the digital compute, by 1959 it was possible to conduct dynamic or full model simulation of the Klein-Goldberger (1955) model. In this sense, this model was not only a pedagogical reference on how to build large scale models or a tool to carry on policy evaluation, but it also became an example of how to use the digital computer in macroeconomics, as well as how to develop a test that would measure model performance in its full form and not just through reduced form equations.

VIII. Full Model Simulations as a Turing Test

Full model simulation could be understood as a sort of Turing test, or as an imitation game (Turing 1950).¹⁸ In order to pass the Turing test, the model should be able to generate data series that are indistinguishable from the data actually observed in the economy. In other words, the model should be able to imitate reality. If the model-generated data produced a pattern that was comparable (although not necessarily equal) to the “real-world” data, then

¹⁸ The contemporary discussion was held neither in terms of conducting a “Turing test” nor of an “imitation game” of course.

Figure 1. Klein-Goldberger's Time Path Consumption Equation

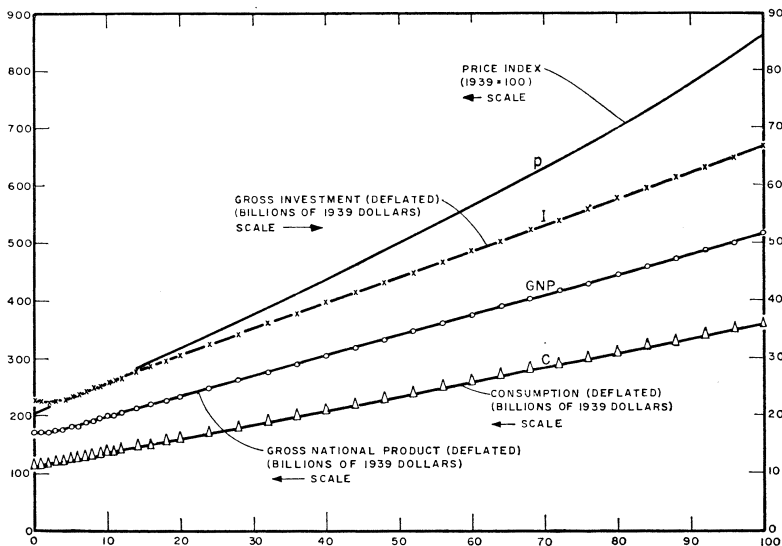


Source: Klein and Goldberger (1955, p. 93).

the model would fulfil this test.

This was, in fact, the objective of the Adelmans' (1959) simulation of the Klein- Goldberger (1955) model. In particular, Adelman and Adelman (1959) examined whether the model “really [offered] an endogenous explanation of a persistent cyclical process [...] whether the system [was] stable when subjected to single exogenous shocks, what oscillations (if any) accompany the return to the equilibrium path, and what is the [response] of the model to repeated external and internal shocks”. (p. 597) To do this, the Adelmans simplified the 22 simultaneous equations model through algebraic substitutions, into a “set of four simultaneous equations in four unknowns” (600). This allowed them to significantly reduce the computational time of the model. The Adelmans performed at least four different simulations in which they asked different questions. The first simulation dealt with the deterministic form of the model and showed that in absence of external shocks, the system was monotonic and essentially linear. Figure 2 presents the results of the first simulation:

Figure 2. *Klein-Goldberger Time Paths (without any shocks)*



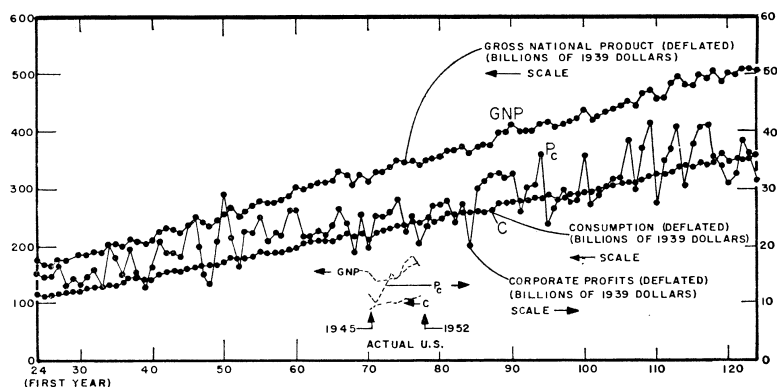
Source: Adelman and Adelman (1959).

The other simulations included different types of shocks. First, they simulated the behaviour of the system when “shocks of type I” were introduced.

This type of shocks consisted in the introduction of a severe and sudden shock in terms of the reduction of government expenditures. As this simulation produced cycles that were “unrealistically small” (p. 609), another type of shocks was introduced.

Shocks of type II transformed the deterministic system into a stochastic system through the introduction of errors into each of the estimated equations. This time (Figure 3), the model appeared to be more realistic since it not only produced three- to four-year cycles, but most importantly, the magnitudes of the cycles “here compare[d] favourably in magnitude with those experienced by the United States economy after World War II” (p. 611).

Figure 3. *Klein-Goldberger Selected Time Paths under Type II Impulses*



Source: Adelman and Adelman (1959).

The importance of the Adelmans’ simulations, especially of those of the stochastic Klein-Goldberger system, resided on its resemblance to the behaviour of the US economy. This resemblance, however, was not only qualitative in nature but also quantitative.

All in all, it would appear that there is a remarkable correspondence between the characteristics of fluctuations generated by the superposition of random shocks upon the Klein-Goldberger system and those of the business cycles which actually occur in the United States economy. The resemblance is not restricted to qualitative

parallelism, but is, indeed, quantitative, in the sense that the duration of the cycle, the relative length of the expansion and contraction phases, and the degree of clustering of peaks and troughs are all in numerical agreement (within the accuracy of measurement) with empirical evidence (Adelman and Adelman 1959, p. 614).

Full model simulations such as that performed by the Adelmans' (1959) allowed macroeconometricians to judge models' performance, taking into account simultaneous equations characteristics that Friedman and Becker (1957) had been missing in Keynesian models. These full model simulations would become increasingly important for the practice of macroeconomic modelling, becoming an important and concrete way in which ideas about dynamics entered this scientific practice.

Conclusions

In a nutshell, Friedman and Becker's (1957) criticism of the "statistical illusion" in judging Keynesian models consisted of two points. The first focused on how to correctly evaluate model performance. According to Friedman and Becker, the performance of the entire macroeconomic model should not be evaluated on the basis of the "apparently" good performance in prediction of the consumption function alone. Implicitly, this point was directed towards the idea that the single equation characteristics are not sufficient to evaluate the performance of the whole model. Instead, to make a sound evaluation of the performance of their models, econometricians should focus on the structural equations characteristics, and undertake whole model simulations. This argument would become common ground in the subsequent years among macroeconometricians, but, at the time, some confusion still existed about the necessity of this criterion. Adelman and Adelman's (1959) simulation of the Klein- Goldberger (1955) model contributed to the establishment of whole model simulation as a criterion of model performance evaluation. Other important contributions in this direction were the further development of the digital computer and the dissemination and further development of the macroeconomic modelling approach.

Friedman and Becker's (1957) second point of criticism was related to the specification of the consumption function itself. In this respect, they insisted that the Keynesian function had to be changed and improved. The important points here are that Friedman and Becker proposed not only two statistical criteria to guide the specification of the new consumption function (the relative error in predicting income from the reduced form equations and the "naive models"), but they also advanced another argument in favour of the permanent income hypothesis proposed by Friedman (1957).

These points, however, have to be understood at two different levels. First, as a claim inscribed within a larger debate between two empirical approaches to macroeconomics, and thus as a claim directed towards the abandonment of the large-scale macroeconometric approach. More importantly to my purpose here is the other level, which interprets Friedman and Becker's (1957) claim as a constructive critique and as precursor of a criterion to evaluate model performance that would become common ground around macroeconometricians: full model simulations or dynamic model simulations.

Acknowledgements

The author would like to thank Emilie Lefèvre, Annie L. Cot, Juan Acosta, Marcel Boumans, Kevin D. Hoover, Pedro Garcia Duarte, the participants of the Albert O. Hirschman seminar of the University Paris 1 Panthéon-Sorbonne, the HOPE Workshop, and two anonymous referees for their valuable comments and remarks on earlier drafts of this paper.

References

- Adelman, I., & Adelman, F. (1959). The Dynamic Properties of the Klein-Goldberger Model. *Econometrica*, 27 (4), 596-625. <https://doi.org/10.2307/1909353>
- Adelman, I. (2007). The Research for The Paper on The Dynamics of the Klein-Goldberger Model. *Journal of Economic and Social Measurement*, 32 (1), 29-33. <https://doi.org/10.3233/JEM-2007-0269>

- Ando, A., & Modigliani, F. (1969). Econometric Analysis of Stabilization Policies. *The American Economic Review*, 59 (2), 296-314. <https://www.jstor.org/stable/1823683>
- Bloor, D. (1976) *Knowledge and Social Imagery*. The University of Chicago Press.
- Bodkin, R.B. (1995). The Discussion among Future Nobel Laurates Becker, Friedman and Klein on Macro Models and Consumption Functions, in 1957 and 1958. In M. Dutta (ed.), *Economics, Econometrics and the LINK* (pp. 45-57). North-Holland Publishing.
- Bodkin, R.G., Klein, L.R., & Marwah, K. (1991). *A History of Macroeconometric Model-Building*. Edward Elgar.
- Boumans, M. (2005). *How Economists Model the World into Numbers*. New York: Routledge.
- Boumans, M. (2016). Milton Friedman and the Cowles Commission. In R. A. Cord & J. D. Hammond (eds.) *Milton Friedman: Contributions to Economics and Public Policy* (pp. 585-604). Oxford University Press.
- Boumans, M., & Dupont-Kieffer, A. (2011). A History of the Histories of Econometrics. *History of Political Economy*, 43 (annual suppl.), 5-31. <https://doi.org/10.1215/00182702-1158781>
- Brown, T. M. (1952). Habit Persistence and Lags in Consumer Behavior. *Econometrica*, 20 (3), 355-71. <https://doi.org/10.2307/1907409>
- Christ, C.F. (1951). A Test of an Econometric Model for the United States, 1921-1947. In NBER (ed.), *Conference on Business Cycles* (pp. 35-130). National Bureau of Economic Research.
- Friedman, M. (1946). Lange on Price Flexibility and Employment: A Methodological Criticism. *The American Economic Review*, 36 (4), 613-631. <https://miltonfriedman.hoover.org/internal/media/dispatcher/214921/full>

- Friedman, M. (1949). The Marshallian Demand Curve. *The Journal of Political Economy*, 27 (6), 463-495. <https://www.jstor.org/stable/1826553>
- Friedman, M. (1951). Comment. In NBER (ed.) *Conference on Business Cycles* (pp. 107-114). National Bureau of Economic Research.
- Friedman, M. (1953). The Methodology of Positive Economics. In M. Friedman (ed.), *Essays in Positive Economics* (pp. 3-43). Chicago University Press.
- Friedman, M. (1955). Leon Walras and His Economic System. *The American Economic Review*, 45 (5), 900-909. <https://www.jstor.org/stable/1821386>
- Friedman, M. (1957). *A Theory of the Consumption Function*. National Bureau of Economic Research.
- Friedman, M., & Becker, G. S. (1957). A Statistical Illusion in Judging Keynesian Models. *Journal of Political Economy*, 65 (1), 64-75. <https://www.jstor.org/stable/1824834>
- Friedman, M., & Becker, G. S. (1958a). The Friedman-Becker Illusion: Reply. *Journal of Political Economy*, 66 (6), 545-547. <https://doi.org/10.1086/258106>
- Friedman, M., & Becker, G. S. (1958b). Reply to Kuh and Johnson. *The Review of Economics and Statistics*, 40 (3), 298.
- Friedman, M. (1958c). The Friedman-Becker Illusion: Supplementary comment. *Journal of Political Economy*, 66 (6), 547-549. <https://doi.org/10.1086/258107>
- Friedman, M., & Meiselman, D. (1963). *The Relative Stability of Monetary Velocity and the Investment Multiplier in the United States, 1897-1958*. New York: Random House.
- Friedman, M., & Schwartz, A.J. (1963). *A Monetary History of the United States, 1867-1960*. National Bureau of Economic Research and Princeton University Press.

- Goldsmith, R. (1955). *A Study of Saving in the United States*. Princeton University Press.
- Haavelmo, T. (1943). The Statistical Implications of a System of Simultaneous Equations. *Econometrica*, 11 (1), 1-12. <https://doi.org/10.2307/1905714>
- Haavelmo, T. (1944). The Probability Approach in Econometrics. *Econometrica*, 12 (annual suppl.), i-vi+1-118. <https://doi.org/10.2307/1906935>
- Haavelmo, T. (1947). Methods of Measuring the Marginal Propensity to Consume *Journal of the American Statistical Association*, 42 (237), 105-122. <https://www.tandfonline.com/doi/abs/10.1080/01621459.1947.10501917>
- Hood, W. C., & Koopmans, T.C. (1953). *Studies in Econometric method*. John Wiley & Sons.
- Hoover, K.D. (1988). *The New Classical Macroeconomics*. Blackwell Publishers Inc.
- Johnston, J. (1958). A Statistical Illusion in Judging Keynesian Models: Comment. *The Review of Economics and Statistics*, 40 (3), 296-298. <https://doi.org/10.2307/1927426>
- Keynes, J. M. (1965) *The General Theory of Employment, Interest, and Money*. Harvest/Harcourt, Inc (original version published in 1936).
- Klein, L. R. (1958). The Friedman-Becker Illusion. *Journal of Political Economy*, 66 (6), 539-545. <https://www.jstor.org/stable/1826994>
- Klein, L. R. (1969). Estimation of Interdependent Systems in Macroeconomics. *Econometrica*, 37 (2), 171-192. https://econpapers.repec.org/article/ecmemetrp/v_3a37_3ay_3a1969_3ai_3a2_3ap_3a171-92.htm
- Klein, L., & Goldberger, A. (1955). *An Econometric Model of the United States, 1929-1952*. North-Holland Publishers.
- Koopmans, T. (1950). *Statistical Inference in Dynamic Economic Models*. John Wiley & Sons.

- Koopmans, T. (1957). *Three Essays on the State of Economic Science*. McGraw-Hill.
- Koyck, L. M. (1954). *Distributed Lags and Investment Analysis*. North-Holland Publishing Co.
- Kuh, E. (1958). A Note on Prediction from Keynesian Models. *The Review of Economics and Statistics*, 40 (3), 294-295. <https://doi.org/10.2307/1927425>
- Lange, O. (1944). *Price Flexibility and Employment*. The Principia Press.
- Marshall, A. (1895). *Principles of Economics, third edition*. Macmillan and Co.
- Modigliani, F. (1949). Fluctuations in the Saving-Income Ratio: A Problem in Economic Forecasting. *Studies of Income and Wealth, XI*. New York: National Bureau of Economic Research.
- Mosak, J.L. (1945). Forecasting Postwar Demands: III. *Econometrica*, 13 (1), 25-53. [https://doi.org/0012-9682\(194501\)13:1<25:FPDI>2.0.CO;2-T](https://doi.org/0012-9682(194501)13:1<25:FPDI>2.0.CO;2-T)
- Pinzón-Fuchs, E. (2016). *Macroeconometric Modeling as a 'Photographic Description of Reality' or as an 'Engine for the Discovery of Concrete Truth'? Friedman and Klein on Statistical Illusions* [CHOPE Working Paper, No. 2016-26]. <https://www.econstor.eu/bitstream/10419/155455/1/chope-wp-2016-26.pdf>
- Pinzón-Fuchs, E. (2017). *Economics as a 'tooled' discipline: Lawrence R. Klein and the making of macroeconometric modeling, 1939-1959* [Doctoral dissertation, PhD thesis, Université Paris 1 Panthéon Sorbonne, Paris]. <https://theses.hal.science/tel-01794108/document>
- Prais, S. J. (1956). *Some Problems in the Econometric Analysis of Company Accounts* [paper presented at the European meeting of the Econometric Society, Aix-en-Provence].
- Singh, B., & Ullah, A. (1976). The Consumption Function: The Permanent Income versus the Habit Persistence Hypothesis. *The Review of*

Economics and Statistics, 58 (1), 96-103. https://econpapers.repec.org/article/tprrestat/v_3a58_3ay_3a1976_3ai_3a1_3ap_3a96-103.htm

Smithies, A. (1945). Forecasting Postwar Demand: I. *Econometrica*, 13 (1), 1-14. [https://doi.org/0012-9682\(194501\)13:1<1:FPDI>2.0.CO;2-7](https://doi.org/0012-9682(194501)13:1<1:FPDI>2.0.CO;2-7)

Stone, R., & Rowe, D. A. (1956). Aggregate Consumption and Investment Functions for the Household Sector Considered in the Light of British Experience. *Nationalökonomisk Tidskrift*, 94 (1-2), 1-32.

Turing, A.M. (1950). Computing Machinery and Intelligence. *Mind*, 59 (236), 433-460. <https://phil415.pbworks.com/f/TuringComputing.pdf>