
Roger E. Backhouse and Beatrice Cherrier

Abstract (110 words)
The FMP model exemplifies the Keynesian models later criticized by Lucas, Sargent and others as conceptually flawed. For economists in the 1960s such models were “big science”, posing organizational as well as theoretical and empirical problems. It was part of an even larger industry in which the messiness for which such models were later criticized was endorsed as providing enabling modelers to be guided by data and as offering the flexibility needed to undertake policy analysis and to analyze the consequences of events. Practices that critics considered fatal weaknesses, such as intercept adjustments or fudging, were what clients were what clients paid for as the macroeconometric modeling industry went private.

Keywords

Acknowledgements
Earlier versions of this paper were presented at workshops at the University of Lille on March 8, 2016, and at the University of Utrecht on April 6, 2017. We are grateful to
workshop participants and, in particular, to Juan Acosta, Kevin Hoover and an anonymous referee for extremely helpful comments.

1. Introduction

The 1970s were a decade of dramatic changes in both macroeconomic performance and macroeconomic theory, which makes it easy to forget what had gone before. As early as 1978, Robert Lucas and Thomas Sargent wrote that they had spent time describing the “halcyon days of Keynesian economics” because “without conscious effort they are difficult to recall today” (Lucas and Sargent, 1978[1981], 295). They described Keynesian economics as involving “a simple set of quantitative relationships between fiscal policy and economic activity generally.” A key point in their description was that Keynesianism could not be understood apart from econometric modeling:

The Keynesian Revolution was, in the form in which it succeeded in the United States, a revolution in method. … [I]f one does not view the revolution in this way, it is impossible to account for some of its most important features: the evolution of macroeconomics into a quantitative, scientific discipline, the development of explicit statistical descriptions of economic behavior, the increasing reliance of government officials on technical economic expertise, and the introduction of the use of mathematical control theory to manage an economy. (Lucas and Sargent 1978[1981], 296; emphasis in the original).

They concluded that Keynesian macroeconomics had failed because it lacked “a sound theoretical or econometric basis” (1978,14), criticizing Keynesian models on theoretical grounds – for a “failure to derive behavioral relationships from any consistently posed

1 The notion that Keynesian models have “failed” or can be considered a “failure” is repeated 17 times.
dynamic optimization problems,” forward-looking expectations and continuous market clearing – and pairing that critique with a long rebuttal of the restrictions used by Keynesians to achieve identification.

Without wishing to pronounce on what constitutes the correct approach to macroeconometric modeling,\(^2\) we argue that acceptance of Lucas and Sargent’s arguments has obscured the goals and practices of Keynesian macroeconometric modelers, and that there is a need for a historical investigation of their epistemological commitments – of what they believed constituted a sound and scientific model. We provide such an investigation through examining the history of one of the main macroeconometric models, the so-called FMP model, named after its three main institutional supports, the Federal Reserve Board, the Massachusetts Institute of Technology and the University of Pennsylvania.\(^3\) Between 1966 and 1970, this model was jointly developed by a team of academic macroeconomists led by Franco Modigliani (MIT) and Albert Ando (Penn), and the staff of the Division of Research and Statistics (DRS) of the Federal Reserve Board, led by Frank DeLeeuw and Edward Gramlich. The resulting multi-equation model was explicitly designed to “say more than existing models about the effects of monetary policy instruments” (De Leeuw and Gramlich 1968, 11). When the original team stopped working on it, the model was transferred to Wharton Associates for maintenance. We claim that, at least during the first decade of its life the FMP model played an important role in the evolution of macroeconomics and exhibited many features common to most large-scale models of the period.

Our history of the model substantiates Lucas and Sargent’s claim that macroeconometric modeling was an integral part of the history of macroeconomics. However, it argues that

---

\(^2\) For a recent debate over this, see the symposium in the *Oxford Review of Economic Policy* 34(1–2), 2018. Some participants, including Olivier Blanchard and Simon Wren-Lewis, argue for approaches that have much in common with the approach we are describing here.

\(^3\) It has also been called the MPS model (where the S denotes the Social Science Research Council) and other names. See Rancan (2018).
Keynesian modelers had epistemic commitments very different from those of Lucas and Sargent. A major reason why Keynesian economics cannot be understood apart from macroeconometric modeling is that there was no linear relationship running from economic theory to empirical models of specific economies: theory and application developed together. The process of macroeconomic modeling was, as Lucas and Sargent alleged, a mess in the sense that the models did not follow rigorously from a consistent theory, estimated using formal procedures that remained stable over time. However, this was not because the creators of the FMP model did not understand what they were doing: it was because they had different goals and criteria. There is indeed a danger that historians let the definition of a satisfactory macroeconomic model provided by Lucas, Sargent or some other “influential” economist, frame objects we are studying. Doing this allows the history of macroeconomics to be written as a succession of well-identifiable models with consistent theoretical foundations and clearly defined estimation strategies. However, history is messy in that boundaries are often blurred and, until we understand economists’ motivations, the procedures followed can seem casual, inconsistent and ever-changing.

How models mediate between the world and economists, and how they and facts travel, being transformed in the process have received substantial attention from historians (see Morgan 2012, Boumans 2015). Our account of the FMP model augments this literature by providing a window into what happen when a model “in the making” is being built and used by several protagonists with different, sometimes antagonistic purposes, tools and values: academic economists, policy makers and businessmen.

Our account shows that it is not even clear that it is possible to speak about “the FMP model.” Instead there was collection of equations whose scope and contours could be adapted to the purpose of model building, and a set of tools (computers and software) and administrative procedures for achieving this efficiently. The problem here, therefore, is not
merely to uncover the compromises reflected in the model, but to track a constantly changing object better described as a collection of models. We therefore examine the different reasons our protagonists had to participate in the FMP project, how these different goals translated into varied simulation practices, tradeoffs between theoretical consistency and forecasting performance, and housekeeping procedures. In other words, we trace the consequences of moving from Keynesian economic science as it existed in the 1950s to the macroeconometrics industry that had developed by the 1970s.

This paper claims that Keynesian economists were aware of both the constant interplay of theoretical and empirical practices that characterized macroeconometrics and the resulting “messiness” of the model. More than this, they endorsed it. Responding to Lucas and Sargent, Ando (1981, 346), spelled out three characteristics of a “sound theoretical and econometric basis” for a model: “(1) all behavioral and other assumptions for deriving the equation to be estimated must be explicitly spelled out. (2) All available evidence from whatever the source, formal or informal, should be marshalled to check the goodness of approximation of all hypotheses at all levels,” and (3) that aggregate hypotheses should be checked again micro data. However, he immediately went on to explain why “it is so difficult to follow this simple principle in building and improving macroeconometric models.” There was a problem with economic theory in that theoretical propositions typically refer to equilibrium states, whereas the real economy is a complex dynamic system which needs to be simplified in order to be analyzed. The need to simplify—to reduce a complex system to a few simple equations—requires “very complex and strong assumptions” (Ando 1981, p. 347). The whole model cannot be tested against time series data alone and, because it was necessary to use different types of evidence, and because there was no statistical theory that could be used to judge whether approximations were adequate, the task of checking
and improving approximations had to be undertaken “informally and implicitly” (Ando 1981, 355).

These informal processes led to the apparent messiness of macroeconometric modeling, but those involved did not perceive this as a failure: it was, rather, messiness by design. In short, not only did Keynesian and New Classical economists have theoretical and empirical disagreements—they also disagreed on which sets of practices would produce a model of “value in guiding policy” (Lucas and Sargent 1978, 2).4

2. A common model for many purposes

Modigliani has claimed that the FMP model was commissioned in 1964 by the Board of Governors, who channeled money through the Social Science Research Council (SSRC) so that the model be developed in universities, thereby reflecting independent judgments about the structure of the economy (Modigliani 2001, 100). His recollections are, however, not entirely consistent with those of the other researchers involved.5 Several economists of the DRS, including Daniel Brill and Frank de Leeuw, had already been involved in the development of a macroeconometric model under the sponsorship of the SSRC, which became known as the Brookings model (Acosta and Pinzon-Fuchs 2018). De Leuw had been in charge of writing up the financial sector, producing a nineteen-equation model which included several policy instruments and transmission channels, from cost-of-capital and wealth effects to credit rationing (see Acosta and Rubin 2018). The Board had declined to support the project financially, in spite of the DRS directors’ attempts to advertise the

---

4 Ando argued that the methods proposed by Lucas and Christopher Sims did not avoid the problems that they said plagued Keynesian macroeconometrics, but involved making arbitrary assumptions that could not be directly verified (see, for example, Ando 1981, 355-6).

5 This paragraph builds on Acosta and Cherrier (2018), who provide a more exhaustive account of how macroeconometrics came to the Federal Reserve Bank. See Rancan (2018) for a different perspective.
“potentialities of econometric methods of projecting the economy’s future performance” (Acosta and Cherrier 2018, 18).

By 1965, Brill, who had by then become DRS director, was making plans for “a comprehensive research project in linkages between monetary policy and the general economy…from Federal reserve actions to spending decisions…among money market variables…and more basic financial variables, such as between bank reserves and the money supply” (Acosta and Cherrier 2018, 19). One reason for developing such model was that the Board had come under intense attacks from Congressmen (notably Wright Patman), some businessmen, and economists ranging from Council of Economic Advisers member James Tobin to monetarists Karl Brunner and Alan Meltzer.⁶ Despite their different views on monetary theory and what constituted sound monetary policy, they agreed that the Board’s decision-making process was “rel[y]ing more on a general faith that virtue pays than on careful empirical and theoretical analysis” (Tobin 1961, 26). The Board thus needed a model that could help them produce the forecasts they were just beginning to use for the meeting of the Federal Open Market Committee (FOMC) and that could show the channels whereby acting on discount rates, borrowed and unborrowed reserves and other policy instruments would affect the economy. Their project was intended as a device for both forecasting and empirical policy analysis.

Building a model that would allow an empirical study of the “complex interaction of forces through which monetary and fiscal policies affect the entire economy,” as Ando would later put it to the Board, was also appealing to academic Keynesian economists.⁷ But their motive was different: they wanted to settle a theoretical controversy. As Gramlich told the National Association of Business Economists in 1969, “[Modigliani and Ando] were spurred on in an attempt to resolve their inconclusive interchange with Friedman-Meiselman in the

---

⁶ Brunner and Meltzer wrote a report for Patman’s Commission on Money and Credit in 1964.
⁷ Ando, “Introduction,” undated but probably 1968, Box RW15 folder ‘notebook,” FMP.
The background to this was the attempt by Milton Friedman, Anna Schwartz and other collaborators to rehabilitate the quantity theory of money, developing the body of thought that Brunner was shortly to call monetarism. As part of this research, Friedman and David Meiselman (1963) produced evidence that the correlation between money and output was higher and more stable than the Keynesian multiplier, implying, so they claimed, that this supported the quantity theory. Ando and Modigliani (1965) responded by arguing that Friedman and Meiselman had failed to specify correctly the variables involved. The ensuing debate, which took a hundred pages of the December 1965 issue of the American Economic Review, centered on estimating two-variable reduced-form equations. Even out-of-sample tests failed to show that one formulation was best, leading to the conclusion that such models could never resolve a theoretical controversy (see Brainard and Cooper 1975, 170; Blaug 1980, 216-17). To do that it was necessary to turn to more detailed structural models, which explains the interest of Ando and Modigliani in building a large-scale, structural macroeconometric model.

As Ando explained shortly afterwards, around 1966 “it was agreed that the two efforts be merged to construct a single econometric model.” Between the Fall of 1966 and the end of 1967, the team specified the theoretical structure of the behavioral equations, made preliminary estimates of the parameters and ran trial simulations of the model. There was a continual interaction between theoretical and empirical work. The builders of the FMP model

---

8 Gramlich, “Complicated and simple approaches to Estimating the role of money on economic activity,” 06/05/1969, Box 1, Gramlich papers. See Rancan (2018) for a detailed account of the monetarist vs Keynesian battle and how it fed into a model war which opposed the Board’s FMP and the Federal Bank of St Louis’s Anderson-Jordan equation.

9 Brunner is usually credited with coining the term “monetarism”, but it was already well established in the literature on Latin America. For example, Seers (1963) used the term to denote an alternative to structuralist policies.

10 Much of Friedman’s work involved establishing correlations and lag lengths, for which Keynesians criticized him on the grounds that correlation and timing did not imply causation.

11 Ando, 1968, Introduction, ibid. The reasons for this decision are unclear. We speculate that Ando and Modigliani might have recognized that the Fed economists had more expertise in the monetary transmission mechanisms, while they had a more refined theory of how financial variables affect the real economy. It might also be that the Fed economists had better access to monetary and financial data needed to estimate the model.
drew on pre-existing economic theory but the process of modeling prompted the development of new theories. For example, Dwight Jaffee and Modigliani (1969) developed a theory of credit rationing, Jaffee also doing extensive research, some of which was with other economists involved in the FMP model, on mortgage markets and housing (see Jaffee 1971, Jaffee and Gramlich, 1973). Rasche (1972) developed a theory about how the performance of the stock market affected consumption, and Modigliani and Robert Shiller (1973) extended the theory of the term structure of interest rates to include expectations of inflation. Ando, Modigliani, Rasche and Stephen Turnovsky (1974), also responding to recent concern with expectations of inflation, developed a theory of how interest affected investment decisions (failure to forecast investment accurately was identified as one of the main problems with the model).

The project thus generated a stream of papers investigating, *inter alia*, various aspects of the transmission mechanism of monetary policy, including the effects of interest rates on housing and plants investment and durable goods consumption, the consequences of credit rationing, and the effects of expectations of future changes in asset prices on the term structure and on the structure of banks’ and households’ portfolios, and Tobin’s $q$. It was also necessary to write software both to do the necessary estimation and to run the simulations. This makes it clear that, even in the first year of its operation, the “FMP model” in fact hosted a multiplicity of models and theoretical viewpoints that were difficult to reconcile, a variety of simulation practices, and diverging ideas of what the model was for and how scientific integrity and forecasting performance should be weighed against another. Macroeconometric modeling was already becoming a messy business.

12 Modigliani Papers, Box CO1(Ando-FRB), untitled, undated statement about end of project
3. Theoretical consistency to messy practice

3.1 A collection of models in constant evolution

During this early period, separate variants of the model were maintained by the academics at MIT and Penn, and by the Fed. It was not until 1969 that these two variants of the model were merged. Even with the software they had developed, merging the models and re-estimating the equations could take two to three months. They reported that, it was not until early 1970 that “for the first time we had something that could be called the FRB-MIT-Penn model.”

The first phase of the model’s life formally concluded at the end of 1970, when the contract between the Federal Reserve and the SSRC was terminated. Modigliani and Ando had taken on the task of providing a written description of the model and arrangements were made for Wharton Econometric Forecasting Associates (WEFA) to maintain and distribute it.

There were several reasons why two sets of models were maintained in the early years. One reason was that whilst the MIT-Penn version of the model would be public, the Fed wished to keep modifications to its version of the model confidential. Brill wrote to Modigliani,

One matter about which I have some concern would be the description of any simulation results of the joint model you might want to publish when our joint effort has been completed. … Obviously it would have to be made clear that these results did not necessarily reflect the views of the Board or the staff. Clearly, we cannot—and would not—try to limit your use of the model in any way. Nonetheless, I can foresee

---

14 WEFA was a non-profit organization established in 1969 by the University of Pennsylvania to handle the many activities associated with the Wharton macroeconometric model.
the possibility of distinct embarrassment to the System from widespread publicity given to strange results of early simulation runs of an untested model.\textsuperscript{15}

Another reason was their different time perspectives. The academic economists wanted a model to adjudicate a long-term controversy but the Fed economists wanted forecasts as soon as possible. Ando acknowledged this in a letter to Brill: “Those of us from the academic side were inclined to postpone the work involved in putting together the model and concentrate on improvements of each equation, but Frank DeLeeuw and Ned Gramlich felt, quite properly, that they must have a functioning system as soon as possible.”\textsuperscript{16}

There was effectively an even wider variety of models in that the Fed economists were running many kinds of simulation, each answering a different type of question. “Diagnostic simulations” were aimed at understanding the characteristics of the model: whole blocks were taken as exogenous, so as to pin down causes and effects in the rest of the system. Software was developed to make it easy to isolate a variable and to substitute its observed values for those predicted by the equations.\textsuperscript{17} “Dynamic simulations” involved feeding into the model forecasts from the previous period into the model for up to 38 quarters, and checking whether the model blew up (it often did) or remained stable and yielded credible estimates for GDP or unemployment (Gramlich 2004). “Stochastic simulations” were carried out by specifying initial conditions, then making out-of-sample forecasts. Policy simulations involved using the multipliers implied by a calibrated version of the model to examine the effects of changes in policy instruments (Gramlich 1997, p. 24). To run these many simulations, the Fed version of the model was ready by November, 1967, while the Penn-MIT version of the model only became operational in the summer of 1968.

\textsuperscript{15} Brill to Modigliani, July 6, 1966. Box 151, folder 1725, SSRC2, cited in Acosta and Cherrier (2018)
\textsuperscript{16} Ando to Brill, 01/10/1969, “Memorandum on administration, Project manager and disbursement of materials, econometric model project,” box RW15 folder “notebook,” Modigliani papers
\textsuperscript{17} Ando to Modigliani, December 8, 1967, p. 2. Modigliani Papers Box CO1 (Ando).
The FMP modelers also realized that, not only did they have to work with several models, but each of them was constantly evolving over time. Initially their view was that they would estimate a specific model—one set of equations—and that this model would then be used for forecasting and policy analysis. However, as they built the model they discovered from their own experience something that other modelers had known—that modeling was an ongoing process, for there were always further modifications to make. In the final report it was made clear that a major part of the model was set of procedures that had been established to manage it. In the period when there were two variants of the model, “housekeeping procedures” had to be devised so as to incorporate the best of both models. The completion of the model involved not just a complete set of equations but the development of “a means of rapidly revising and re-estimating it.” A report on maintenance of the model stated that “it is expected of any econometric model that, as time passes and new data accumulate, some equations in the model begin to drift. In some cases, this is because of the changes in institutional arrangements, in others, it is because of the subtle shifts in the empirical definitions of variables, and in still others, due to unsatisfactory specifications of the theory underlying the specific equation.” The model was something that needed to be managed, with some watching over its performance over the years, a task that Rasche was willing to undertake. The model had become something that evolved.

3.2 Different tradeoffs between theoretical consistency and forecasting performance

Yet another reason for maintaining several versions of the model was different approaches to theoretical specification. The Fed economists seemed more concerned with institutional details, as attested by the many monetary control variables in De Leeuw’s equations: non-

---

18 For example, by the end of the century, the Bank of England had concluded that “economic models should not be thought of as fixed in form or content, and that model development is a continual process” (Bank of England 2000, p. 365).
19 Ando to Hickman, Undated memorandum, “Maintenance of MPS model”, Modigliani Papers, Box CO1 (FRB).
borrowed reserves, reserve ratios against demand and time deposits, the discount rate and ceiling rates on bank time deposits. His theory was informed by a close knowledge of Fed operations. Modigliani and Ando adopted a different approach in which the integrity of key theoretical insights had to be preserved. Acosta and Rubin (2018) explain that the number of policy variables in DeLeeuw’s equations was subsequently drastically reduced to a decision concerning free reserves. The need for competing theoretical specifications within the same macroeconomic modeling project was well understood and accepted. Ando wrote, “None of us holds the view that there should be only one model. It would indeed be unhealthy if there were no honest differences among us as to what are the best specifications of some of the sectors of the model, and when such differences do exist, we should maintain alternative formulation until such time as performances of two formulations can be thoroughly compared.”

The flexibility with which the theoretical structure was handled when it did not perform well during simulation shines through Ando and Modigliani’s correspondence. To understand how messiness was embedded in the design of the model, it is worth quoting one of their exchanges at length. After a series of simulations which had not blown up in the Fall of 1967, Ando wrote:

This does not mean that we do not have some major problems to work on beginning in January. First of all, the rent equation seems to be completely unsatisfactory, and we had to leave it out for our simulation. I think the consumer durables equation must be improved substantially. Of course, there are sectors which we have left out: labor market and wages and prices, and the mortgage market. On the labor market, it may be that manhours equation and the hours per man equation which Richard [Sutch] has given us, together with the participation equation that we put together last summer,

---

will be satisfactory, but we have not tried yet. That is the next step in our simulation experiments. On wages, [Alfred] Tella tried to carry on where I left off last summer, and ran into all sorts of trouble. I still think that if we worked sufficiently hard, something can be salvaged, but it must be worked on further. On the general price level, you must have got the equation that [Richard] Sutch reluctantly gave us, and as he says himself, it is not a satisfactory equation. Of course, there is the question of relative prices, which we have not touched….

On the question of the mortgage market, I went down to Washington this Tuesday….and talked it over with [Gordon] Sparks and Frank [de Leeuw]. Sparks now has the demand equation for mortgages, and a set of equations for the supply of mortgages by savings institutions. They look at least promising, and we can use them as a starting point. On the other hand, he still does not have a completely satisfactory equation for housing starts, so the housing sector will have to be worked on quite a bit … [Harold] Shapiro has done no work whatsoever on the consumption sector, so that too will have to be taken over by us. This sounds rather formidable, but I don’t think the task will be impossible…21

How the equations were handled reflected economists’ tradeoffs between analytical consistency and forecasting performance. Such tradeoffs were inevitable because predictive success required modifications that had either had no theoretical justification or even ran counter to economic theory. This can be illustrated by discussions over the consumption sector of the model. Remarks made by De Leeuw in March 1968, when he wrote that dynamic simulations were improved if current income was dropped from the consumption equation illustrate how the equations had to be modified in order to reduce simulation errors: “We get a considerable reduction in dynamic simulation errors if we change the total

consumption equation by reducing the current income weight and increasing the lagged income weight [...] We get a slight further reduction of simulation error if we change the consumption allocation equations so as to reduce the importance of current income and increase the importance of total consumption,” he explained, explaining that the consequence of the revision would be that “the multipliers will build up more gradually than in our previous policy simulations, and also that the government expenditure multiplier will exceed the tax multiplier.”

The exchange with Modigliani also shows that Fed and academic economists were willing to make different kinds of trade-off between theoretical consistency and predictive power. Modigliani, who needed to solve a theoretical debate, objected to DeLeeuw’s treatment of consumption, writing “I am surprised to find that in these equations you have dropped completely current income. Originally this variable had been introduced to account for investment of transient income in durables. This still seems a reasonable hypothesis.”

As explained by Acosta and Cherrier (2018), the practice at the DRS was to further blend the “mechanical forecasts” produced by the model with “judgments” provided by those in-house “business economists” who understood and anticipated the “feel and tone” of several sectors of the economy. Chairman Martin remained extremely skeptical toward macroeconometric modeling, so that theoretical consistency had no appeal whatsoever to the Board.

Diverging goals therefore nurtured conflicting model adjustments, even though it is unclear how much FMP was used as a policy aid in the end. By the time the model was in full operation, Arthur Burns had replaced Martin as chairman. Though a highly skilled economist and former Chair of the CEA, he proved as skeptical toward macroeconometric models as his predecessor, and his decisions were largely driven by political pressures (Ferrell 2010). Several economists recall that the Bluebook, the document produced to inform discussion at

22 De Leeuw to Ando, March 6, 1968. Modigliani Papers, Box CO1 (FRB).
meeting of the Federal Open Market Committee (FOMC) and intended to present 3 alternative monetary scenarios, was closely overseen by Burns. The staff was, according to former Fed staff member Arnold Kling (undated, 26), not free to come up with whatever forecast it thought most probable. Board member Sherman Maisel, DRS econometrician James Pierce and DRS advisor Peter Keir all stated that Burns would ask that option B, the middle-range one FOMC members were supposed to converge toward, be a write-up of his preferred policy rather than a reflection of independent forecasts (Acosta and Cherrier 2018). Such an approach was justifiable given that even supporters of the model recognized its limitations. For example, Lyle Gramley, a senior economist at the Fed who later became a Governor, supported the model but would not trust it as a forecasting device. Ando admitted to the SSRC that he was not optimistic about the short-run forecasting ability of the model, and Gramlich was also confessed that “our model has left something to be desired on this score [forecasting].” Behind diverging practices and trade-offs there were different conceptions of what counted as “science,” and of what kind of science was most useful for forecasting and policy purpose. The disagreement with other groups of economists, in particular monetarists and New classical were even larger.

4. From Keynesian Science to macroeconometric industry

4.1 Macroeconometric modeling as “big science”

By the standards of MIT, which achieved its post-war eminence in large part through being the site of massive natural science and engineering laboratories, building the FMP model was a small project, but from the perspective of economics in the 1960s, it was “big

24 Modigliani Papers, CO1(FRB), Gramley to Ando, April 29, 1968 shows his support for the model; Gramlich 2004 documents Gramley’s scepticism about its forecasts.
science.” Given the computing technology of the day and the need to draw on a wide range of expertise, such models had to be created by teams of researchers, and it posed bureaucratic as well as technical challenges. The construction of the model involved not simply Ando, Modigliani, De Leeuw and Gramlich themselves, but a large team of economists, including many MIT graduate students: Larry Meyer was in charge of the housing sector (modeling how equity and housing values are affected by monetary policy). As was mentioned earlier, Jaffee worked with Modigliani on the effect of credit-rationing on housing, de Menil handled the wage equation, with a focus on the effect of unions on wages, while Charles Bischoff provided a model of plant and equipment investment, and Sparks wrote the demand equation for mortgages. Ando concentrated on estimates of fiscal multipliers whilst Modigliani researched how money influenced wages and links between expectations and interest rates with students Sutch and Schiller. At the Fed, Enid Miller, Helen Popkin, Tella and Peter Tinsley, worked on the banking financial sector and transmission mechanisms, in particular portfolio adjustment. Responsibilities for data compilation, coding, running simulations were also shared between academics and the Fed, with Rasche playing a key role. By the standards of much economic research of the day, therefore, this was “big science.”

Initially, a major reason for revision of the model was the process whereby it was constructed. It seems that graduate students picked a topic, then worked in relative isolation for months, gathering their own data, surveying the literature on the behavior of banks, firms, unions consumers or investors before sending back a block of equations. Because these blocks each had different structure, characteristics and properties, various methods were used to estimate them. Even if each block was a good fit to the data on which it was estimated, there could be no assurance that it would produce sensible results when combined with other blocks estimated on different data, something that could be ascertained only through simulations. Later on, changes in the economic environment, notably significant increases in
the inflation rate, necessitated reworking many sectors of the model, and so a new batch of researchers in various universities reworked the housing, consumption, financial and investment blocks in 1969-1973.

As has been already been documented (Hoover 2012), Keynesian macroeconomists did not ignore the need for microfoundations, but problems such as these meant that ensuring consistency was a major headache. There was no assurance that the separate parts of the model would fit together, despite Modigliani and Ando being in overall charge of the project. When a dynamic simulation was conducted looking 8 years (32 quarters) ahead and the model did not explode, keeping reasonably close to the observed data, this was considered a major achievement.26 This problem was not, of course, confined to the FMP model. For example, a report from 1969 on the Brookings model notes that “When the original large-scale system was first planned and constructed, there was no assurance that the separate parts would fit together in a consistent whole” (McCarthy 1992). Consistency involved more than standardization of the theory and econometric procedures used but extended to the “housekeeping procedures” that enabled the modelers to keep track of alternative equations and how they worked together, and to quickly revise and re-estimate the model. These included building large database capabilities with easy access and update procedures, common packages, such as AUTO-ECON or TROLL, the residual-check procedure aimed at automatizing code-accuracy checks (McCarthy 1992).

A further problem was that computational capabilities and simulation procedures were, by modern standards, undeveloped, though it is not clear how far this shaped the models and their results. 1960s reports are filled with computer breakdowns and coding nightmares. For example, in 1967, Ando lamented, “the reason for the long delay is […] that the University of Pennsylvania computer facilities have completely broken down since the middle of October

during the process of conversion to a [IBM] 360 system, and until four days ago, we had to commute to Brookings in Washington to get any work done.”

The complexity of the task of performing the simulations was no doubt a source of frustration, yet these computational constraints could have stimulated the creativity of the economists working on the model. The creative solutions had to become “material,” for instance by being embedded in software. One example is the 1967 simulation package (SIM) developed by Morris Norman for econometric model simulation which allowed modelers to solve a system of equations through Gauss-Seidel iteration faster than through block procedure (McCarthy 1992, 398).

The FMP model was not just “big science” but was part of a wider “industry” of macroeconometric modeling, for it was one of many models created at this time. This was significant because different models were not developed in isolation from each other. There were frequent consultations between modelers, in part because of Klein’s presence at the University of Pennsylvania and because of the role of the Social Science Research Council in funding several models. When the SSRC and NBER planned a conference to analyze the performance of different models, Klein urged Ando to participate on the grounds that otherwise all the models [OBE, Wharton and Brookings] would be “products of his work, one way or another.”

Not that the FRB model was completely independent: it has been claimed that it “grew out of the Brookings project” and that material on the public sector was taken from the Wharton model (Bodkin, Klein and Marwah 1991, p. 108). As the model was developed, Ando discussed it with Klein. The scheme developed by Norman and Raasche for processing data and managing the FMP model was “adopted almost bodily” by those operating the Wharton and Brookings models, and the eventual ongoing maintenance of the model by Wharton, which had its own model, illustrates how close different models could

be. Such practices were a long way from the vision of scholars working in isolation except in so far as they read and responded to newly published material and engaged in common-room discussions with colleagues in the same institution. The scale of this industry and the extent to which it overlapped with macroeconomics more generally make it an important part of the history of macroeconomics.

All in all, the model was simply too complicated for any analytical solution to be possible. The model as a whole was therefore, like all comparable models, a mess in that the interaction of different sectors might produce behavior that was clearly ridiculous. It therefore had to be calibrated in ways that had no clear connection with the underlying economic theory. For example, intercept adjustments were commonly used. Such adjustments could be taken so far that it was possible for a model to be “transformed so many times since its inception that it [became] a model almost devoid of theory, with equations altered and dummy variables added wherever necessary to maximize the model's ability to produce a control solution that accurately tracks GNP during the sample period” (Gordon 1972, p. 298). This was a description of the Brookings model, where the process had gone further than with the FMP model, but the apparent messiness of both models arose, at least in part, from the challenges posed by building such a large model given the available economic theory and the available computing technology and was made worse by the different goals motivating the projects.

4.2 Rejecting or endorsing messiness

This messiness was identified and challenged by opponents. Brunner (1973, 929-931) considered the fundamental problem to be a confusion of what he called “pragmatic” and

31 Unlike Lucas, Gordon (1972, 298) believed that the Brookings model was devoid of theory but contrasted it with “the MIT-FRB and Wharton Models, which are both supported by a considerable body of theoretical and econometric literature.”
“cognitive” claims. So much “pragmatic numerology” was needed to get the models to fit the data that being able to produce good forecasts (the pragmatic claim) did not constitute evidence for the theories on which the models were based or for their policy analysis (the cognitive claim). In his view, then, the model could not settle the disagreement with Friedman and Meiselman over monetary policy. The Federal Reserve Bank of St. Louis, with which Brunner and Meltzer were eventually affiliated, had developed a much smaller model inspired by monetarist ideas (see Rancan 2018), and agreeing on what constituted a convincing empirical test proved difficult. The objective was to ascertain whether changes in short-term market interest rates were independent of changes in the stock of interest-bearing government debt held by the public, Modigliani having claimed that they were, and Brunner and Meltzer that they were not. Presumably intending to avoid the inconclusive results of earlier tests of Keynesian and monetarist theories, Meltzer proposed a series of tests and how the key variable, government debt, would be defined.\textsuperscript{32} Replying for Modigliani, who was then on his way to Italy, Rasche agreed to three of the four tests, and made suggestions about who should compile some of the necessary data. It is not clear whether the tests were actually carried out.\textsuperscript{33}

The exchanges between Brunner and Meltzer, and the MPS team were not on the kind of statistical test that should be conducted and on the data that should be used. They concerned the precise kind of prediction a “good” model was supposed to yield and on what tests it was possible to carry out. For instance, The St Louis economists proposed that their models should be able to “predict the changes in short-term market interest rates from the 1st quarter 1962 through 4\textsuperscript{th} quarter 1966, using the regression equation for your hypothesis and data through the 4\textsuperscript{th} quarter of 1961.”\textsuperscript{34} The FMP team rejected one of Meltzer’s proposed tests on

\textsuperscript{32} Meltzer to Modigliani, 17 July 1967, Modigliani Papers, Box CO1 (Ando-FRB).
\textsuperscript{33} Rasche to Meltzer, September 7, 1967, Modigliani Papers, Box CO1 (Ando-FRB).
\textsuperscript{34} Meltzer to Modigliani, 17 July 1967, Modigliani Papers, Box CO1 (Ando-FRB).
the ground that the Federal Reserve did not have some of the data. However, it was seen as a hindrance that there was no agreed statistical test to compare alternative model specifications.

The FMP model was constantly compared with models developed by the Brookings Institution, the Wharton School, the Federal Reserve Bank of St. Louis and other organizations, and public comparisons were carried through conferences and volumes sponsored by the NBER (Hickman, 1972; Fromm and Klein, 1976). This led the FMP team to establish a partnership with statisticians. “We must develop a more systematic procedure for choosing among the alternative specifications of the model than the ones that we have at our disposal. Arnold Zellner of the University of Chicago has been working on this problem with us, and Phoebus Dhrymes and I have just obtained a National Science Foundation grant to work on this problem,” Ando wrote in 1968.35 Zellner was already employed as a consultant statistician and was assisting with the conduct of simulations.

One of the practices criticized by Lucas (1976), in his influential critique of econometric modeling, was “intercept adjustments,” sometimes known as “fudging” or “add factors” (Fair 1986). If a model produced apparently systematic errors (if, for example, there was a run of positive residuals) this could be corrected by revising the intercept in the corresponding equation downwards by an equivalent amount, improving the accuracy of the next forecast. Such adjustments became increasingly necessary after the structural changes and inflationary pressure that followed the 1973 oil crisis, when models systematically under-predicted inflation, leading economists such as Lucas to conclude that they had irretrievably broken down. He pointed out that such methods produced better forecasts, but any improvement reflected the judgment of the forecasters—it did not validate the models: “the unquestioned success of the forecasters should not be construed as evidence for the soundness or reliability

of the structure proposed in that theory” (Lucas 1976, 23, emphasis in original).

The fact that a model was producing systematic errors that needed to be corrected by intercept adjustments meant that it must be misconceived. Lucas elaborated on this point using the examples of consumption and investment expenditure and the Phillips curve. The need for intercept adjustments and other modifications to estimated equations could arise from failing to model expectations as rational. Though Lucas confined his examples to single-equation models, his arguments would clearly apply, a fortiori, to large-scale models: he described “the adaptations necessary for simultaneous equations” as being “too well known to require comment” (Lucas 1976, p. 26)

However, “fudging,” the practice Lucas condemned so harshly, became a major reason why businessmen and other clients would pay to access the forecasts provided by the FMP and other macroeconometric models. Ray Fair later noted that analyses of the Wharton and Office of Business Economics (OBE) models showed that ex-ante forecasts from model builders (with fudge or add factors) were more accurate than the ex-post forecasts of the models (with actual data): that “the use of actual rather than guessed values of the exogenous variables decreased the accuracy of the forecasts” (Fair 1986, 1985-6). In the late sixties and early seventies, several forecasting companies were created to host major macroeconometric models: while Wharton WEFA was a non-profit, Data Resources International (DRI), established by Harvard economist Otto Eckstein, and Chase Econometrics, established by Klein’s colleague Michael Evans quickly became profitable. According to Kling, and consistent with Fair’s conclusion, cited above, the hundreds of fudge factors added to large-

---

36 Further evidence was that the process of re-fitting equations did not seem to be converging on more precise parameter estimates.
37 The OBE was in the US Department of Commerce; in 1972 it was reorganized into the Bureau of Economic Analysis.
38 An unsigned article from Time magazine, dated Monday, Jun 25, 1979, titled “Business: Flash and a Touch of Brash,” reports that Chase was by then a “$100 million-a-year business.” That same year, the Harvard student newspaper, The Crimson reported that DRI was sold to McGraw-Hill for over $100 million (https://www.thecrimson.com/article/1979/7/17/mcgraw-hill-inc-plans-to-buy/).
scale models were precisely what clients were paying for when buying forecasts from these companies. These forecasts were “providing us with the judgment of [Otto] Eckstein, [Michael] Evans and Adams as embedded in the add factors, and these judgments are more important to most of their customers than are the models themselves” (Kling, undated, 26). 39

What Lucas perceived as bad science was hailed by the models’ customer as a sign of flexibility and a guarantee that the model would take into account recent business climate shifts. Different communities of economists had different attitudes toward theoretical consistency.

4. Conclusion

For many economists, Lucas’s critique was decisive. It became widely believed that, because people could learn from their mistakes, economists should adopt the theory of rational expectations, and that alternative methods of forecasting and policy analysis needed to be developed. In the 1980s and 1990s, therefore, a wide range of new modeling strategies were pursued, from atheoretical time-series analysis to the development of DSGE (dynamic, stochastic general equilibrium) models, many of them immune to the Lucas critique. However, although there was a move towards developing simpler, theoretically coherent models, large-scale Keynesian models such as the FMP were never abandoned completely. Small, highly aggregated models, even if they embodied rigorous theory and rational expectations, could neither provide the policy analysis needed by the Fed and other organizations, nor the predictions and scenarios for which businessmen were willing to pay (Webb 1999). The FMP model was kept going until 1995, during which time it was not merely maintained but it was developed. It had initially contained around 65-70 behavioral

39 There is insufficient evidence conclusively to identify Adams. It would seem most likely to be F. Gerard Adams, who worked on the Wharton model (Bodkin, Klein and Marwah 1991, p. 127 and passim). However, the 1970 American Economic Association membership directory also lists a Robert W. Adams whose 1951 PhD dissertation, supervised by Paul Samuelson, was on forecasting.
equations, but by 1985 the number of such equations had grown to 124 and there were 197 exogenous variables (Rasche and Shapiro 1968, 123; Brayton and Mauskopf 1985, 170). The model might be inconsistent and need continual revision, and its results might need continually to be checked against other models and other types of evidence, but it was considered indispensable for macroeconomic policy analysis by enough people to justify this work (Brayton, Laubach and Reifschneider, 2014). It was eventually replaced by a model that allowed the Federal Reserve Board to choose between a menu of expectations equations, showing that, at least in one central bank, Lucasian types of modeling did not replace Keynesian ones, but were treated as additions to the available toolbox.

Lucas and his followers might criticize Keynesian macroeconometric modeling for its lack of theoretical consistency but theoretical consistency could be achieved only at the price of ignoring many of the demands that caused Keynesian economists to adopt the methods they did. The messiness of the mode may or may not have been a fatal flaw, as Lucas claimed, but it was the result of economists engaging in a large-scale scientific project that posed technical and organizational problems that needed to be solved. Maintaining a variety of sub-models within the projects, taking blocks of equations in and out for simulations, fudging intercepts, adding new constraints or behavioral equations that were not micro-founded was a conscious strategy to yield better predictions and useful conditional forecasts and policy analysis. Where critics saw messiness, macroeconometric modelers saw flexibility, a characteristic of models which is increasingly being advocated in post-crisis macroeconomics. In 2017, in advocating the development of specific models for “policy analysis,” Vitor Constancio (2017), Vice-president of the European Central Bank, noted that “we constantly update our beliefs on the key economic mechanisms that are necessary to fit the data,” concluding that “the model should be reasonably flexible.” Olivier Blanchard (2018, 51) likewise argued that theory models and policy models should be developed
separately: “policy modelers should accept the fact that equations that truly fit the data can have only a loose theoretical justification.” In a surprising turn, he acknowledged that “in that, the early macroeconomic models had it right: the permanent income theory, the life-cycle theory, and the Q theory provided guidance for the specification of consumption and investment behaviour, but the data then determined the final specification.” Simon Wren-Lewis (2018) has even argued that, had traditional structural econometric models, such as the ones analyzed here, been maintained alongside models with more rigorous microfoundations, economists would have been in a better position to respond to the challenges posed by the Great Recession. The revival of such attitudes does not mean that the FMP modelers and their contemporaries were correct, but it does suggest that we should make a serious attempt to take seriously their motivations and the rationale they saw for the methods they employed.

Reference

Gramlich Papers – Ed Gramlich Papers, Bentley Historical library, University of Michigan.

Modigliani Papers - Franco Modigliani Papers, Rubenstein Library, Duke University.


Rancan, Antonella, 2018. The Keynesian and Monetarist debate on structural versus reduced form models: the Fed MIT Penn and the St. Louis models, details to be confirmed.


