February 2016

# Art, Science, and the Impact of RBC Models in Economic History

by

Alexander J. Field
Department of Economics
Santa Clara University
afield@scu.edu

ABSTRACT

Thirty years have now elapsed since Larry Summers judged that "real business cycle models of the type urged on us by Prescott have nothing to do with the business cycle phenomena observed in the United States or other capitalist economies." In 2004 Kydland and Prescott won the Nobel Prize in economics, and their work has encouraged new methods and objectives for macroeconomic research and new standards for explanation and model evaluation, inspiring some economists and economic historians and provoking others. This paper considers the impact of RBC theory in macroeconomics and macroeconomic history, and asks, in light of what has transpired, whether Summers' evaluation should be challenged, modified, or overturned. The paper concludes that economic historians have been justified in their limited attention to real business cycle theory. We should remain aware of the unusual evaluative methods it employs and the damage the spread of RBC empirical tools has done to our understanding of the macroeconomy. Although the more radical claims of RBC proponents have been widely rejected, the diffusion of favored statistical tools, in particular the HP filter, has damaged our ability to think about and formulate policy, in particular by corrupting our understanding of output gaps, and how we measure them.

## Introduction

In 1982 Finn Kydland and Edward Prescott published “Time to Build and Aggregate Fluctuations”, an event generally recognized as marking the birth of real business cycle theory (the terms were first used by John Long and Charles Plosser in 1983). Kydland and Prescott endorsed an increasingly influential view that the behavior of macroeconomic aggregates had to be modelled as the collective outcome of the decisions of hundreds of thousands, even millions, of individual optimizing agents. Building on Lucas’s policy ineffectiveness postulate and their earlier “Rules Rather than Discretion” paper (1977), they challenged the utility of systems of equations macroeconometric models and the ways in which such models were then used to predict the effect of individual policy changes. Most distinctively, they insisted that business cycles (now business cycle phenomena) had to be accounted for in terms of the same supply side factors used in the Solow model to explain long run economic growth. Combining these elements, Kydland and Prescott launched a frontal assault on the entire IS-LM neo-Keynesian framework that had dominated teaching research, teaching, and policy work since the early 1950s. The 1974-75 recession had exposed apparent vulnerabilities in that approach. Both unemployment and inflation soared, calling into question the short run Philips curve as it was then understood. In advancing their agenda, Kydland and Prescott exploited the intellectual disarray induced by the unusual economic history of the decade of the 1970s.

The principal counterweight to neo-Keynesianism in the late 1960s and 1970s was monetarism; the two most well-known proponents of these approaches were Paul Samuelson and Milton Friedman, who aired their differences in academic venues as well as dueling Newsweek columns. The two frameworks were, however, more similar than their spokespersons were prepared to acknowledge at the time. Each located the principal source of business fluctuations in aggregate demand phenomena – changes in the growth rate of nominal income. Although they had different perspectives on how stable was monetary velocity, Neo-Keynesianism and monetarism formed part of the same macroeconomics language group, a fact that became much more apparent in the light of the RBC challenge. If neo-Keynesians spoke French and monetarists spoke Italian, RBC modelers spoke the equivalent of Hungarian or Finnish.

In 1986, four years after Kydland and Prescott's paper, Larry Summers published an assessment of their contribution to research in macroeconomics. His blunt summative assessment: "…real business cycle models of the type urged on us by Prescott have nothing to do with the business cycle phenomena observed in the United States or other capitalist economies" (1986, p. 24).

Kydland and Prescott went on to receive the 2004 Nobel Prize in economics. Their work encouraged new methods and objectives for macroeconomic research and new standards for explanation and model evaluation, inspiring some economists and economic historians and provoking others. Although a number of the more radical RBC claims are now broadly rejected, their modelling strategy and favored empirical and evaluative methods have, in contrast, achieved substantial penetration within mainstream macroeconomic research. Three decades have now elapsed since Summers advanced his judgment, more than three decades since Kydland and Prescott published. In the light of what has transpired, should Summers' evaluation be challenged, modified, or overturned?

Macroeconomics is often described as a nonexperimental science. There may sometimes be reasons to doubt whether it is truly a science – although it certainly has aspirations in that direction. If we pursue those aspirations, we must rely, in our quest to improve our knowledge of how the world works, on observational data and the "experiments" provided by history. In attempting to understand the remarkable economic history of the first decades of the twenty first century, there has consequently been great interest in plumbing the histories of the Great Depression, as well as cyclical phenomena both preceding and following it: the 1907 crisis that led six years later to the establishment of the Fed, the 1974-75 downturn that witnessed the birth of the concept of stagflation, the 1980/1982 downturns resulting from the attempts to bring inflation under control by restricting the growth of monetary aggregates, as well as many other intervening if more moderate booms and busts. This has been true for other countries as well. Following the financial crisis of 2007-08, the disciplinary boundaries between macroeconomists and economic historians writing macroeconomic history blurred even more than they had in the past.

Macroeconomic history had always been, or should have been, a central concern for economic historians. Macroeconomic historians brought perspectives, empirical knowledge, and understanding of data sources of particular interest to the larger profession. Especially during the years surrounding the financial crisis and ensuing recession, there was considerable intellectual ferment, as scholars reacted and responded to the extraordinary episode in economic history then playing out. Then, as economies recovered, and memories faded, there was a gradual falling off of interest in what might be learned from history, probably beginning around 2010. Some things had changed, perhaps permanently, but much had not, as scholars returned, more and more, to business as usual. By 2016 the practice of macroeconomics appeared less altered than many had hoped or expected in 2008 or 2009.

Still, the cross fertilization of macroeconomics and macroeconomic history operated positively in both directions. Policymakers turned to history for guidance as they devised responses to the crisis and Great Recession. At the same time, the experience of a potentially catastrophic world economic meltdown led economic historians to examine previous macroeconomic episodes in a new light. Recent and not so recent history thus provide a context within which to assess whether, or how, or in what ways the RBC initiative has impacted the study of macroeconomics and macroeconomic history. The boundaries between these are not distinct. There is no authoritative decree that work on data prior to a certain date is macroeconomic history, yet subsequent to it is to be consider macroeconomics per se. All macroeconomic data has been generated in the past, and those working on more recent data write the first drafts of macroeconomic history.

In addressing this question we will touch on how much insight RBC models provided in diagnosing and treating the consequences of the financial crisis and great recession before considering how much help they may have given us in understanding the 1930s and other periods. The stakes were and are very large. The catastrophe in terms of cumulative output loss that was the great depression is of course well known. But that associated with the great recession may, by some measures, end up in the same order of magnitude. By 2016 it was clear that either a very substantial output gap remained or the consequence of eight years of negative gaps had bent downwards the trajectory of potential from where it had been projected by the Congressional Budget Office in 2007 (Field, 2015). In either case, the damaging consequences could not be gainsaid.

Let us begin by stipulating that RBC models offered little help in devising policy responses to the stunning collapse of hours and labor force participation in 2008 and 2009. This is perhaps not surprising, since the RBC framework implied that a main tools of policy makers – monetary policy – could have only minimal effects on such variables.[1] The idea that negative supply shocks caused the 2007-09 recession seemed at the time as implausible to most macroeconomists as had similar claims made previously about the Great Depression to economic historians. If RBC models proved of little assistance in addressing recent challenges, the question is whether they can or have enriched our understanding of prior history.

**The Impact of RBC models in Economic History**

A quick overview of the literature leads to the conclusion that the RBC impact in the subfield of economic history has been minimal. A JSTOR search for the words 'real business cycle' in The Journal of Economic History for the years 1980 through 2010 reveals 1 full text article reference, a contribution addressing the procyclicality of TFP between 1890 and 2004 (Field, 2010). References in four book reviews also turn up. A similar search within the Economic History Review returns hits in two book reviews and none within articles. A search in Science Direct within Explorations in Economic History identifies one article using RBC methods. On the other hand, if one does a similar search in the Journal of Political Economy or the Quarterly Journal of Economics the hits are almost too numerous to count.

How do we interpret this? Certainly there has been research on the Great Depression and other downturns influenced by RBC methodology; work by Cole and Ohanian (1999, 2000, 2002, 2004) and the collection of contributions assembled in Kehoe and Prescott (2007) are perhaps most notable. One might also mention Bordo, Erceg, and Evans (2000). Why has

[1] Fiscal policy (government spending shocks) can affect current output and employment in RBC models, but not, as in the traditional IS/LM model, via multiplier mechanisms operating in an economy with slack, where the size of the multiplier depends in part on the degree of monetary accommodation. An output gap is ruled out in an RBC model, which also predicts that consumption will decline and investment might rise, the opposite of what empirical evidence (and IS/LM theory) would predict (see Galí, López-Salido, and Vallés , 1992). In any case, no RBC proponent recommended either monetary stimulus or a massive infrastructure or other federal spending program as a response to the economic downturn.

virtually none of this admittedly small body of research been published in the flagship journals of economic history? It's possible, of course, that because such research has been so important and has attracted such widespread interest, there have been opportunities, even for those who self-identified as economic historians, to publish in general interest journals such as the AER or the JPE (this has been true in other areas of macroeconomic history). Or it might be that economic historians and journal editors have been skeptical of RBC models, like Summers, questioning their value, and tending to reject papers in this tradition. Christy Romer said of the work of Cole and Ohanian, "I think it is wonderful when anyone works on the Great Depression. But honestly, I think some of this research is a giant step backwards (Parker, 2007, p. 135). RBC modelers tackling topics in macroeconomic history may have anticipated some of this, and thus, given the choice between a possibly sympathetic general interest journal and an unsympathetic field journal, submitted to the former.

Still, even if this is right, we must ask what lies behind the negative evaluations of the contributions of RBC influenced research among prominent macroeconomists and economic historians. There are several commonly cited reasons. Reaching a judgment similar to Summers', Greg Mankiw emphasized two: "…real business cycle theory does not provide an empirically plausible explanation of economic fluctuations. Both its reliance on large technological disturbances as the primary source of economic fluctuations and its reliance on the intertemporal substitution of leisure to explain changes in employment are fundamental weaknesses" (1989, p. 79).

**Technology Shocks as the Cause of Business Cycles: The Impulse Mechanism**

Technology shocks are the most common of the supply side disturbances appealed to within RBC models to generate business cycle phenomena. An HP filter is applied to total factor productivity data, separating trend from cycle. The deviations from trend are identified as technology shocks and are the impulses in RBC models generating fluctuations in output and other macro variables.[2] The shocks alter the marginal products of labor and capital and thus the returns to working and saving/investing. In response, people vary their labor supply and consumption behavior, propagating the impulses, and generating business cycle phenomena in output, hours, consumption, and saving. These changes impact the flow of capital accumulation, meaning that the phenomena may have persistence even when the initial impulse is transitory. This channel, is nevertheless, relatively weak in RBC models; most persistence is the consequence of persistence in the shocks themselves.

A major problem with this explanatory agenda, one repeatedly emphasized by economic historians and others, is that technology shocks are both positive and persistent, and if negative they are generally transitory. Aside from disasters like the destruction of the library in Alexandria, or perhaps the collapse of the Roman empire, our scientific and engineering knowledge doesn't typically regress. If the deviations from trend TFP truly measure technology shocks, it would be quite surprising if, against a background of positive trend TFP growth, negative deviations were sufficiently large to cause the level of total factor productivity to fall.

It might be plausible to imagine that the differential rate of arrival of positive shocks was responsible for variations in the rate of growth of TFP and actual output. Certainly it is widely accepted that TFP growth plays an important role in the longer run growth of potential output. In the RBC universe, there is no distinction between actual and potential output, so it is natural that technology shocks should be seen as significant in accounting for both the short and long run

[2] In Kydland and Prescott the impulse series is computer generated, constrained to have the same standard deviation and autoregressive structure as the Solow residuals. Subsequent authors experimented with the use of the actual residuals themselves.

behavior of macroeconomic aggregates. In summarizing Kydland and Prescott's contributions, the Nobel committee explained that "periods of temporarily low output growth … could simply follow from temporarily slow improvements in production technologies….recessions (could be) caused by lower-than-average technology growth leading workers to work fewer hours and consumers to invest less" (2004, pp. 3, 14).

None of this language, however. addresses the possibility (and reality) of declines in output during most recessions. It is hard to believe, in the absence of plausible narratives, that negative productivity shocks could account for big declines in output, such as occurred between the years 1929 and 1933, or for that matter, between 1981 and 1982. Calomiris and Hanes (1995) emphasize this, and go on to question whether TFP deviations should be interpreted as truly reflecting technology shocks. They argue that since important innovations usually take many quarters, indeed years to diffuse, the arrival of one innovation would be unlikely to "increase aggregate productivity by more than a trivial amount from one year to another" (1995, pp. 69-70). If business cycle expansions were typically the consequence of positive productivity shocks, we would, moreover, expect them to be accompanied by decelerations in inflation. Similarly, if recessions were the result of negative shocks (or at least a slower rate of arrival of positive shocks), we would expect them to be accompanied by acceleration in inflation. With the exception of the unusual recession of 1974-75, it is quite rare for such patterns to be observed: typically we see the reverse.

RBC proponents have an answer to the objection that output declines in most recessions. They broaden the universe of supply shocks to include legal, tax, or regulatory changes. Such changes are almost universally considered by those appealing to them to be negative in their effects on productivity; most RBC proponents exhibit great skepticism about the ability of government to benefit its citizenry. Aside from military spending, and the provision of legal and judicial systems to enforce claims to private property, protect persons, and insure that contractual obligations are either met or, in the case of breach, adequate compensation is paid, RBC proponents tend to believe strongly in limited government: governments that govern least govern best.

There are numerous historical example in which political disruptions plausibly cause output actually to decline. For example, most estimates are that the Soviet Union did not reattain the level of output enjoyed by Russia in 1913 until 1926. That said, for a country like the United States which, since the Civil War, has experienced only moderate political disruption and minimal wartime destruction, regulatory changes stemming from legislation or judicial decision are unlikely to be able to account for such large downturns as occurred between 1929 and 1933. Negative law or regulatory changes could help explain slower long term trend growth. But if they are to account for cycles in which output declines and then revives, such shocks must be temporary, and large relative to the positive trend influence of progress in scientific and technical knowledge. They would probably need to occur during periods in which such knowledge was growing relatively slowly or not at all (Stadler, 1994, p. 1771), not the case during the depression years (Field, 2011). Finally, the negative shocks would need to be large enough to overwhelm the possible positive effects on output of an increase in the labor force due to demographic factors and/or a rise in the capital stock. It's conceivable that reduced labor productivity (output per hour) and lower hours per unit of the population due to intertemporal substitution of leisure could overwhelm the positive influences of growing scientific and technical knowledge, population, and physical capital stock. But it would be a steep climb.

The numerous obstacles confronted by a posited negative "technology shock" in actually lowering output and employment help explain why Cole and Ohanian's accounts of the Great Depression, which emphasize the supposed high wage policy in manufacturing encouraged by President Hoover at a conference in November of 1929, and subsequently the National Industrial Recovery Act and other New Deal policies, has been viewed skeptically by most macroeconomic historians. The explanation of the "slow" growth of output between 1933 and 1937 as due to cartelization and lax antitrust enforcement has likewise been doubted. Output growth was in fact extraordinarily rapid under Roosevelt over those years. To be sure, growth did not succeed in closing the output gap at the local peak in 1937: there remained obstacles to recovery in some sectors, especially construction (Field, 1992). But the recovery was, nevertheless, very fast. From a rhetorical standpoint Cole and Ohanian have not been able to weave together a compelling narrative explaining onset, depth, duration, and recovery. For scholars with knowledge of the data and the history, their modeling and empirical work has not proved persuasive. Reasons for this include the nature of the RBC explanatory strategy as well as how its "success" is judged.

The baseline RBC model limits the impulse mechanism to technology shocks. A fundamental objection to this maintained hypothesis is that the strong procyclicality of total factor productivity, an empirical regularity about which there is widespread agreement, is much more plausibly accounted for as almost entirely endogenous, induced perhaps by some labor hoarding but most importantly by the inability to deaccession physical capital (and avoid its holding costs) during recessions (see Field, 2010).[3] That TFP deviations from trend are due to

[3] User costs, which capital holders must bear in bad times as well as good, are governed principally by the real rate of interest and the rate of depreciation. Depreciation flows on the entire capital stock are, contrary to widespread belief, largely invariant to low utilization. This point of view is at variance with standard analyses, which treat machinery and equipment as representative capital goods (they are not) and assert that since recessions are times

technology shocks, narrowly or broadly defined, is simply asserted. Macroeconomic historians are particularly sensitive to broad assumptions or assertions about the long run course of scientific and technical change or the causes of short run fluctuations in TFP not backed by plausible empirical or historical analysis.

**Intertemporal Substitution: The Propagation Mechanism**

The second factor alluded to by Mankiw is the role of intertemporal substitution of leisure, which is critical to propagating impulses in RBC models. Consumers do not just consume based on current income receipts as in a simple Keynesian model, but are sophisticated life-cycle optimizers. In fact they are more than life cycle optimizers. The utility function used by Kydland and Prescott is infinite horizon, suggesting that decision units be thought of as persisting family units with strong dynastic and bequest motives. These households optimize not only over consumption/saving, but also over how much labor to supply in different periods.

Assuming diminishing marginal utility, maximizing the discounted present value of utility over multiple periods implies consumption smoothing. Optimizing agents may therefore increase saving (and thus capital accumulation) in the face of a positive shock.[4] If supply shocks of

when capital lies relatively idle, they must also be times when depreciation flows are low (see King and Rebelo, 1998, p. 980). Mechanical equipment and perhaps some other subsets of equipment subject to friction or heat deterioration may experience less wear and tear when idle. Depreciation due to fashion or technological obsolescence, if applicable, will continue. Machines for which lower utilization means lower depreciation comprise only a portion of the equipment stock. The vast bulk of the net stock consists of structures. Whether a building is full or empty, whether its rate of throughput is high or low, its roof will continue to wear out, and exterior paint will continue to oxidize at about the same rate (Field, 1985, 2010). In some cases, particularly where structures or factories are unoccupied, low utilization may actually increase depreciation. Aside from vandalism, the interiors of shuttered facilities may suffer from inadequate control of temperature, dust, or humidity. Idle machinery may suffer from lack of lubrication. The use for example of electricity input to proxy for variations in capital service flow is thus questionable. In criticizing work which addresses "measurement error" in TFP residuals by making adjustments to capital input based on variable utilization, Prescott (2016, p. 19) actually signs on to much of this analysis of depreciation, although not the larger point about endogeneity that is being made.

[4] It is also conceivable, since preferences are one of the two key primitives in the theory (the other being 'technology'), that an economy might be subject to preference shocks which affected labor supply and/or saving

whatever form make it more attractive to work now than later, then optimizing agents may shift labor supply to the current period. The connection between these two margins for optimization is complex. We must first consider what's involved in intertemporal substitution. One way to motivate the propagation mechanism is to think about tire chain installers on Interstate 80 between the Bay Area and ski areas in the Lake Tahoe region. A heavy snowstorm will cause tens, perhaps hundreds of these individuals to drop whatever else they are doing, put on their yellow rain gear, and head out to where the temperature is dropping to 33 degrees and the driving rain is turning to slush and snow. Clearly these individuals are engaged in some kind of intertemporal substitution of leisure. Similarly, as Bob Hall has observed, most economic activity tends to be bunched during daylight hours. The smaller the region, the sector, and the time period considered, the more compelling are these kinds of narratives. The question is whether they can be taken as plausible metaphors for what happens over extended periods at the level of the national economy. As Paul Krugman has observed, the RBC explanation of the drop in employment during the depression is that hundreds of thousands of people decided to take an extended vacation because a bad technology draw worsened the relative attractiveness of working in 1930, 1931, 1932, 1933, and indeed up through 1941 (this of course is somewhat inconsistent with the suggestion that the problem was that wages were maintained at too high a level during these years).

The mechanics of why, within a standard general equilibrium RBC model, work and leisure are, in response to a technology shock, reallocated over time will not be transparent to an Econ 1 student (or an ordinary human). Labor supply (work/leisure) decisions are affected by

behavior, and thus output and employment in different periods. In that case the preference changes would be considered impulse mechanisms.

both income and substitution effects, operating in different directions (this is also true for the saving/consumption decision). Thus, increased voluntary labor supply could be the direct effect of a positive or a negative supply shock on wages, and the same is true for decreased supply. Most of the microeconomic evidence on labor supply suggests that within the range within which real wages have tended to fluctuate, income and substitution effects roughly balance in the aggregate, leading to an approximately vertical aggregate labor supply schedule. Again, the relative strength of income and substitution effects varies by age, gender, and other demographic group. But overall, it's a wash. As Miles Kimball and Mathew Shapiro have written, "One of the best documented regularities in economics is that—when they affect all members of a household proportionately—large permanent difference in the real wage induce at most modest differences in the quantity of labor supplied by a household. This is true across countries, across households, and across time" (2008, p. 1). Macroeconomic evidence in advanced economies shows little trend in hours per capita in the face of increases in wages over time.

It is sometimes suggested that RBC proponents simply ignore income effects. In fact, the policy preferences of many imply the view that income effects dominate among the poor and substitution effects among the wealthy, leading to a "growth agenda" emphasizing higher taxes net of subsidies for the poor and lower taxes for the wealthy. For the most part, however, RBC models, do not incorporate heterogeneous agents; all households (and all firms) are assumed identical. MaCurdy (1981) provides a detailed theoretical and empirical treatment of labor supply response within the context of a life cycle model similar to that used in RBC models. His estimate of the elasticity of labor supply for prime age males working continuously is .15. RBC modelers must however insist on a much higher intertemporal substitution of leisure (labor supply elasticity) in response to a productivity shock because they need it to motivate the coincidence of large changes in employment over the business cycle as voluntary in the face of relatively small changes over the business cycle in real wages. They dismiss MaCurdy's estimates based on the argument that hours or work are not continuously variable; the appropriate margin for aggregate analysis is in or out of the labor force.

On the theoretical front, RBC models typically employ log linear utility functions, which have zero *static* wage elasticity: at a point in time the income and substitution effects due to a change in wages exactly cancel. This would seem to make it impossible for positive supply shocks to induce higher contemporaneous labor supply due to a change in wages now and/or at any time in the future. The channel from positive supply shocks to increased contemporaneous labor supply and output must instead run not directly from changes in the expected time profile of wages but through the effect of the shocks on wealth (permanent income) interacting with consumption smoothing. An increase in current period labor supply can be thought of as a response to the temporarily higher interest rates consistent with the transition to a new (and higher) capital-labor ratio which results in the allocation of a higher proportion of output to saving and thus capital accumulation.[5] People will want to enjoy their higher wealth through some combination of higher consumption and more leisure. In the short run, however, a desire to spread some of the consumption increment to the future leads to increased saving (capital accumulation). As part of the solution to the general equilibrium optimization problem, the representative agent finds it desirable to provide more labor supply in the current period. Thus there will be an association between higher wages resulting from the supply shock and increased labor supply (and output).

[5] "Thus in the near term real interest rates rise, which induces intertemporal substitution of current for future work effort" (Plosser, 1989, p. 60).

All of this may seem introspectively questionable. RBC proponents insist that, in order to avoid the Lucas critique, macroeconomic models must be built up from the dynamic optimizing decisions of self-interested economic agents. Not all economists, however, agree, and among these economic historians probably figure disproportionately. Mankiw doesn't include this objection in his critique since the principle that microfoundations are essential has been almost uniformly accepted by macroeconomic theorists, including New Keynesians. This has been the price paid in order to reintroduce monetary non-neutrality within the context of models constructed according to RBC principles.

But we may well ask why microfounded macromodels are inherently better. It cannot be that they are more realistic than those lacking them. The baseline RBC model has one actor (a representative agent) standing in for every household in the economy. All agents in the economy have the same tastes and face the same budget constraints. A modeler seriously committed to greater realism and possibly better prediction would probably want, for example, to exploit the fact that higher income individuals have lower marginal propensities to consume than do lower income individuals. The use of a representative agent rules this out. The very serious aggregation problems associated with adding up the decisions of tens of millions of heterogeneous agents are swept under the rug. Conditions for exact aggregation are stringent and unlikely to be realized in the real world (Kirman 1992). Once one allows for heterogeneity, any equation specifying the behavior of a representative agent may well respond quite differently to shocks than will the summation of the responses of individual agents.

Consider a modified Keynesian aggregate consumption function from a large scale macro model from the mid-1960s. One might argue, from a philosophical perspective, that the specification is deficient because all consumption decisions are ultimately made by individuals or households. One might say it would be both more realistic and provide better predictions if one developed an individual consumption function for every household. Such an endeavor, were it practical, would provide real microfoundations for the treatment of consumption in a macro model. But in opting for a representative agent to "stand in" for all individual households, RBC modelers make no effort in this direction. Microfoundations seems to mean nothing more than that the model contains a multiperiod or infinite horizon utility function, with arguments consumption and leisure, that the representative agent maximizes subject to constraint. No matter that all the individuals are assumed identical. The insistence on models of this type reflects an aesthetic preference as much as anything that can be defended on empirical grounds.

If the standard is how realistic are its assumptions, and how reliable are its predictions, the "microfounded" models contained in RBC theory, implemented using a representative agent, do not necessarily offer an advantage over models specifying relationships between macroeconomic aggregates. Those may be subject to the Lucas critique, but how important is this empirically, in comparison with the violence to reality done by the representative agent? Hartley, Hoover, and Salyer argue that in applying the "formal mathematics of microeconomics", RBC modelers "provide the simulacrum of microfoundations, not the genuine article" (1997, p. 39).

Whether the "genuine article" is worth aspiring to is a separate question. On pragmatic grounds it would be impractical to develop a separate consumption function for every household. And it is hardly realistic to imagine that any actual human is capable of the kind of intertemporal optimization of both consumption saving and labor/leisure presumed by these models, nor is it likely to be true that all households share the bequest motive and dynastic impulses reflected in the infinite horizon model.

Of course Milton Friedman advised us back in 1953 to stop worrying about the realism of assumptions. In that sense he was a supporter of the mantra that "all models are false", which has been echoed by Prescott. But Friedman also stressed the importance of judging an economic model by how well it could explain and, ideally, predict.

The basics of RBC methodolology are somewhat unusual from a scientific standpoint. The procedure is to construct an artificial economy, feed it impulse shocks similar to those experienced in the real world, and compare some features of the model's output with some features of actual macroeconomic time series. Building blocks include a representative agent who faces capital inherited from the past, a Cobb-Douglass production function, and a log linear utility function defined over consumption and leisure. Budget constraints insure that total output equals total income equals total expenditure, mimicking the national income accounting identities. Modelers select what they believe to be plausible parameter values including labor's share, the discount rate, and the rate of depreciation on capital. Technology shocks are extracted as the deviations from a TFP trend calculated using an HP filter. The model is "solved" numerically by inputting technology shocks similar to those extracted and having the model generate time series for output, hours, consumption, investment, the capital stock, etc. Standard deviations and co-movements of the generated series are then compared with those of data from the actual economy. A high degree of correspondence is treated as evidence for the success of the model, with modelers expressing surprise and engaging in self-congratulation on the 'surprisingly good fit.'

**Calibration**

"Calibration" involves tweaking parameters in pursuit of an even better correspondence. This iterative process, which involves the comparison of model output with the second moments and comovements of actual macro series leads proponents to describe their inquiry as both scientific and empirical. The facts that the standard deviation and autocorrelation of technology shocks are extracted from actual TFP data, that the model parameters are sometimes estimated econometrically, and that the artificial data are, as described, benchmarked against actual data are used to justify characterizing the research as empirical. But, as Milton Friedman observed in a 1997 letter in the Journal of Economic Perspectives, "There is a world of difference between mimicking and explaining, between 'can or may' and 'does' (1997, p. 210).

Science can tolerate, indeed it depends upon models whose assumptions are in many ways unrealistic. It has difficulty progressing, however, with models whose proponents can't specify results that would lead them to modify their priors, and where evaluative methods can't discriminate among quite different processes that might have produced the same or similar results. Although Friedman was at pains to emphasize that models need not have realistic assumptions, there is a presumption that models based on more realistic assumptions, those that do better at capturing the essential features of the processes in question, are more likely to do better on empirical tests. Research that employs a more traditional methodology contains a feedback loop: A model that performs poorly empirically may lead the researcher to reconsider some of its assumptions.

In contrast, within an RBC model, the assumption that technology shocks are the sole cause of business cycles, which is a really important statement about the world if true, will not be rejected within the confines of RBC methodology. Parameters may be tweaked in search of a better correspondence between characteristics of model output and actual data, but the maintained hypotheses will not be modified. Yet RBC proponents are quite unclear about whether they see themselves engaged in estimation, hypothesis testing, or something else entirely. The point of the research initiative seems to be to provide apparently scientific confirmation of positions arrived at beforehand, in particular the integrated complex of claims that technology shocks are the sole cause of business cycle phenomena, markets always clear, actual always equals potential output, and all fluctuations in employment are voluntary. As Friedman, noted, there is a world of difference between demonstrating that a model based on these principles could approximate features of the data and the conclusion that the exercise has told us much of anything about whether the world actually operates according to them. It should be said that New Keynesian models, which may start with different assumptions but nevertheless proceed using RBC evaluative methods, are subject to a similar critique.

Even if we could agree that the exercise of comparing second moments and comovements has bearing on the validity of the model, proponents typically eschew formal empirical tests of the degrees of correspondence. They simply eyeball the results, pronounce themselves satisfied, and move on. Hartley, Hoover, and Salyer describe this rather casual approach to data as "aesthetic R squared" (1997, p. 42). It is true that the choice of parameters is often justified by appeal to empirical data. But there is no mechanism in these efforts that would allow us to judge whether this a good or bad model of the world according to widely accepted canons of scientific method. Trying to formalize the correspondence or goodness of fit tests is an interesting exercise (Watson, 1993) but doesn't get to the heart of this problem.

The best empirical science allows us to learn something about the real world that we didn't previously know. A commonly accepted (though sometimes criticized) canon of scientific methodology is that a hypothesis or model does not qualify as scientific if it is impossible to conceive of data that would "disprove" it. Are RBC models tested with data? On the one hand RBC proponents appeal to mimicry, judged by a seat of the pants test, as evidence of how good the model is. On the other hand, Prescott has argued that the validity of or confidence one has in these models can't "be resolved by computing how well the economy mimics historical data… the degree of confidence in the answer depends on the confidence that is placed in the economic theory being used" (1991, p. 171). Which is it? This seems to be a formula better suited to religious than scientific inquiry.

The appeal to the degree of confidence one has in the economic theory being used suggests a Bayesian approach. But scientific inquiry extremely becomes difficult to carry forward as the degree of prior confidence in one's model approaches 100 percent. If one holds very strong priors, one will be unlikely to change one's mind in response to new empirical evidence. Might one still be interested in looking at data? Yes, but only insofar as it might help point towards improperly measured variables. In other words, discrepancies between model predictions and data would lead one to question the data, not the model. Prescott's discussion of "Theory Ahead of Business Cycle Measurement" (1986) suggests thinking along these lines.

Is data relevant in evaluating these models? Elsewhere, Prescott writes: "The models constructed within this theoretical framework are necessarily highly abstract. Consequently, they are necessarily false, and statistical hypothesis testing will reject them" (Prescott, 1986, p. 10). What can it possibly mean to say that "highly abstract" models are necessarily false, other than to preclude or express lack of interest in traditional means for judging their validity? Models of planetary motion are highly abstract. Does that mean they are necessarily false? Those models make predictions which can then be tested against astronomical observations. If predictions were at wide variance from observational data, we would be inclined to discard or modify the model, or, we if we had strong priors on its validity, explore the possibilities of measurement error being the cause of the discrepancy.

In his Nobel lecture, Finn Kydland defends the calibration process – adjusting model parameters so that the artificial model economy generates data exhibiting a higher degree of correspondence, based on a comparison of second moments and comovement, with actual macroeconomic time series. He compares the models to measuring devices, which need to be calibrated. Like a thermometer, he suggests, the model should give the correct answer to questions whose answer we already know. But how is a DSGE model like a thermometer? What is the analog of the cold or hot water the model is "measuring?" The Nobel prize committee described calibration as "a simple but informative form of estimation when confronting new models with data" (Royal Swedish Academy of Sciences, 2004, p. 14). What can this possibly mean? Elsewhere the academy writes, "Clearly, calibration is a simple form of estimation, since the model parameters are chosen in a well-specified algorithm to fit a subset of the overall data; in this case, the estimation is based on microeconomic and (long-run) macroeconomic data. However, the method was very practical." (2004, p. 17). Why is the word 'however' used? Again, what can these passages mean?

In their Nobel lectures, both Kydland and Prescott go out of their way to emphasize the quantitative and scientific character of their work. Prescott talks about contributing to a "quantitative scientific discipline" and suggests that in trying to extend Lucas's work on the policy ineffectiveness postulate, he wanted a technique for evaluating the success of different policy rules. He was frustrated and ultimately dissatisfied with the large econometric models of the 1960s and 1970s which were used to evaluate the consequences of an individual policy action.

Both are aggressive in touting the scientific benefits of the methodology they advance. In the video of his Nobel talk, Kydland praises the essays in the Kehoe/Prescott volume (2007) as demonstrating that the proposition that RBC models can't be used to study such large macroeconomic fluctuations as the Great Depression is "bullsh**". In discussing the role of the stock market, Prescott expresses "surprise" at how well his model performs (most variation in stock market values across countries is simply the consequence of higher or lower tax rates on capital), and is dismissive of animal spirits and "behavioral mumbo-jumbo." It would be churlish to suggest that the surprise seems somewhat disingenuous. Objectively speaking, in spite of all the talk about data, the methods used have the effect of supporting prior beliefs. Whether they were consciously developed for this purposes is a separate question. But it is hard to find instances in RBC research where the inquiry has lead a researcher to change his or her mind about a fundamental assumption based on the relationship of model output to data.

Perhaps the most devastating critique of RBC/DSGE methods was published in 2009 by Emi Nakamura. Nakamura shows that "in a model with the same basic structure as the bare bones RBC model, monetary, cost-push or preference shocks are equally successful at explaining the behavior of macroeconomic variables. Thus, the empirical success of the RBC model with respect to standard RBC evaluation techniques arises from the basic form of the dynamic stochastic general equilibrium model, not from the specific role of the productivity shock" (p. 739). In other words, RBC methods are incapable of discriminating among hypotheses that different types of aggregate demand or aggregate supply shocks are the main cause of business cycles. If, within the confines of the models, these widely different types of shocks can generate similar degrees of correspondence between second moments and comovements, the "standard RBC evaluation techniques" can't provide relevant information germane to which of them may be important in causing business cycles. Nakamura's conclusions reinforce arguments made earlier by Hansen and Heckman (1996).

The calibration process therefore appears to be something of a distractor. Initial results are, without much discussion, interpreted as evidence in favor of the validity of the model. The focus immediately shifts to calibration, aimed at smoothing out rough edges and tidying up some odds and ends, creating what looks to be an interaction between theory and data in the best traditions of scientific discourse. It has the effect of diverting attention from the troubling reality that demonstrating rough correspondence of standard deviations and comovements of model output and actual data doesn't provide evidence of model validity, since many different types of shocks are capable of generating similar correspondence.

**Mechanical Dogs**

One might say that RBC models are to the functioning of the macroeconomy as Sony's AIBO, introduced in 1999, is to a real dog. Imagine a lot more R and D poured into that robotic canine, and then measuring how well AIBO mimicked the metrics and comovements of metrics for a live dog. Would this tell us anything about whether the robot's internal electromechanical devices were realistic approximations of the neurobiological mechanisms of a live dog? Would it give us any guidance as to whether we had a "good" model of the dog? I would argue that for many purposes, the answer is no. Place AIBO in a sealed chamber, evacuate the air, and there is likely to be no deterioration in function. Don't try this with a real canine, however: the out of sample prediction will fail.

The RBC revolution was, first and foremost, intended to be about method. Prescott made this very clear in his Nobel lecture. This is important in considering where the initiative failed and where it succeeded within the larger discipline. Many of the extreme claims of RBC proponents have been rejected. These include the view that the principal causes of business cycle phenomena are short run fluctuations in the rate of TFP growth, and the view that the aggregate economy is characterized by high intertemporal substitution of leisure. New Keynesian models resurrect a role for nominal shocks in these models by reintroducing wage and price rigidities. But these exercises don't provide more evidence that the NK view is correct, any more than Kydland and Prescott's initiative provided evidence that technology shocks are the sole cause of business cycle phenomena.

Much of what non-applied economists do is often denigrated as the attempt to show that what works in theory also works in the real world. Are RBC models subject to this critique? Perhaps not, because in spite of protestations to the contrary, the models are not really confronted with data. There is more interest in aesthetics than there is in empirics. Not many minds have been changed; the example of Narayana Kacherlakota an exception that proves the rule. The RBC initiative has, nevertheless, left a negative legacy that goes beyond the low payoff consumption of so much intellectual firepower devoted to elaborating, refining, and critiquing it.

In ways probably unanticipated by Kydland and Prescott when launching their impulse, and largely unappreciated by those who unintentionally propagated it, RBC proponents have ended up doing some very serious damage to our ability to understand and treat cyclical variation in a way that effectively balances policy makers' twin mandates of low inflation and high employment. It has done so by fundamentally corrupting our understanding of output gaps, and how we calculate them.

Even though RBC proponents have no interest in calculating output gaps, the concept – the difference between actual and potential output – is critical to economic policy making. To measure such a gap we need measures of both actual and potential output. Actual output is what the Bureau of Economic Analysis provides us. Potential output – current estimates and historical estimates going back to 1950 are provided by the Congressional Budget Office – reflects estimates of the highest level of output that can be produced without so stimulating the economy that we experience an acceleration of inflation. At potential, the unemployment rate will be the NAIRU – the nonaccelerating inflation rate of unemployment.

**The Concept of an Output Gap: A Brief History**

The concept of the output gap was born of two concurrent and interrelated developments, the construction of national income and product accounting by Simon Kuznets and the publication of the General Theory by John Maynard Keynes. Prior to the General Theory there was no reason to draw a distinction between actual and potential output, and their difference – an output gap – because there was no theoretically coherent explanation of how output might be persistently below potential (this is the state of affairs to which Lucas, Kydland, and Prescott would like us to return). Keynes provided a compelling account of why and how output might stabilize at a point where there was considerable slack in the economy – involuntarily unemployed labor and unused physical capital capacity, and might remain there due to a persisting deficiency of aggregate demand.

These two developments occurred during the 1930s, encouraged and inspired by the plummeting output and employment levels associated with the worldwide depression. The idea of an output gap, as the difference between actual and potential, underlay Kuznets' estimates of how large a military force and how much war material the US would be able to produce during the war, subject to maintaining prewar consumption levels (Field, 2011).

As of the start of the war in 1941, however, the terminology was not yet widely used. JSTOR scans of the journal literature identifies 1945 as the year in which the concept explodes in print frequency. As the end of the war came in sight, a major concern and preoccupation of economists was the prospect of a large output gap following war demobilization, that is, a concern that the slack which had characterized the entire decade of the 1930s would reappear with the withdrawal of the fiscal and monetary stimulus associated with the war. Economists now used the terms actual and potential output, and output gap (their difference) freely, and without need of further explication. Within the space of less than a decade, the terms had became part of an economist's standard toolkit.

We can see this in Richard Musgrave's article about what kind of deficit spending might be needed to close the gap: "Suppose that with a given federal budget over-all income falls substantially short of the potential output at full employment. What adjustment in the budget can be made to raise income to the full employment level?" (1945, p. 387). And we see it in Everett Hagen's survey of ten estimates of the "Postwar Output of the United States at Full Employment", which begins by emphasizing the importance of "estimating the nation's potential postwar output" (1945, p.45), and in Louis Bean's comment on it (1945).

Already it was apparent that there were slightly different challenges involved in calculating a current or retrospective output gap, which involves comparing actual or past GNP with an estimate of potential, and forecasting the future trajectory of potential. Hagen describes how the forecasting project involved using demographic estimates to project potential hours (combining forecasts of the growth of the labor force, frictional or turnover unemployment, and hours per week), and estimates of the trajectory of output per hour (labor productivity). In his article Hagen showed how differences among these ten estimates could be reduced to differences in forecasts of one or more of these components. In 1947 Thomas Schelling, in an article comparing Harrod's and Domar's growth theories, talks routinely about the "potential (full employment level) of income" (1947, p. 865).

The degree to which the concept of an output gap had become commonplace in postwar discussion is underlined in a landmark 1950 AEA report, “The Problem of Economic Stability”, which was intended to be read by the general public. Future Nobelists Milton Friedman and Paul Samuelson were two of the five authors. They wrote, inter alia, “The desirability of full employment is obvious: unemployment means waste of potential output and hardship for the unemployed and their dependents” (1950, p. 506). Although Friedman expressed dissenting views regarding the final section of the report on “Market Policy” he apparently had no problems with the rest. Indeed, potential output was simply what he would subsequently refer to as natural output.

During the Eisenhower years, references declined somewhat, but in the 1960s, with Kennedy’s promise to get the American economy moving again, they picked up again. In 1962 Arthur Okun identified a robust regularity in economics data that came to be known as Okun’s law, a strong relationship between changes in the unemployment rate in percentage terms and changes in the output gap. The inflation of the 1970s, widely attributed to negative supply shocks associated with OPEC’s restriction on the worldwide production of oil, led to some disarray as the Philips curve had to be reconfigured as the inflation augmented Philips curve. Building on the work of Edmund Phelps and Milton Friedman, potential output was now defined theoretically as the highest level of output the economy could sustain without being so stimulated that it experienced an acceleration of inflation.

The intellectual cycle came full circle in the early 1980s, as RBC proponents pressed once and for all to put the genie back in the bottle: to destroy the idea of an output gap as a useful concept, and along with it most of the corpus of Keynesian economics that had helped birth it.

The goal was to return economics to the pre-Keynesian consensus, one reason RBC models are considered to be part of New Classical economics.

RBC proponents did not succeed in eliminating the concept of the output gap from the vocabulary of policy makers. But they did succeed in changing how it is estimated retrospectively, and in doing so weakened the consensus as to its meaning and how it is defined.

The standard means of forecasting the trajectory of potential has not changed greatly since the end of the second world war. Building on growth accounting methodology, the CBO forecasts the growth of quality adjusted hours, based on demographic and other variables, the growth of physical capital, and the growth of total factor productivity. Weighting the first two variables by their share in national income and adding TFP growth generates an estimated time series for potential. In identifying a trajectory for potential output, we are asserting that given sufficient aggregate demand, the economy could perform at these levels without experiencing an acceleration of inflation. One could also, as did Hagen, combine forecasts of the growth of hours and labor productivity.

The dominant procedures used retrospectively to calculate potential output and its growth path have, however, changed. Prior to the RBC revolution the principal method can be described as trend through peaks. Data on actual GNP or GDP are examined – this could be at annual or higher frequencies, and we make attempts to identify years in which output and employment were high, unemployment was low, but in which there is no evidence of an acceleration of inflation. We posit that during those benchmark years the economy was at potential or close to it and approaching it from below. Assume two such "peaks" have been identified. On log linear paper, draw a straight line connecting the peaks and possibly extending before and after them for a number of years. The output gap can be identified as the difference between actual output and the estimate of the trajectory of potential.[6] The method requires careful examination and consideration of data and the use of judgment, and it imposes the requirement of a constant continuously compounded rate of growth for as many years as the line is drawn. Deeply rooted in NBER traditions it nevertheless appears to many moderns as seriously flawed.

Partly because of its apparently primitive nature – redolent of the 1950s and 1960s – this approach is almost never used today. Instead, a sophisticated computer technique features in almost all macroeconomic research – including that conducted by many who do not share the more extreme views of RBC advocates who promoted it. The method involves the use of the Hodrick-Prescott (or similar) filter, into which are fed data on actual output along with a smoothing parameter, lambda. We are interested in outputting a time series on the trend of GDP. A simple OLS regression on logged values of output would produce a coefficient which, multiplied by a time dummy, would generate a series showing trend output. But it would grow at a constant rate. The HP filter allows the trend growth rate to vary continuously over time (thus apparently overcoming one of the objections to the trends through peaks approach) by penalizing changes in the growth rate of the trend. The higher is lambda, the higher is that penalty. At lambda increases, we approach our OLS trend estimating procedure, since as one approaches a constant growth rate one eliminates second differences (changes in the rate of change).

There have been many criticisms of the HP filter, such as the fact that if one uses the latest data following a downturn, one will almost certainly conclude that the year or years immediately preceding the downturn were above potential. But this is not what is truly pernicious. *The more*

[6] Note: prospectively, our forecast of TFP growth governs our estimates of potential, whereas retrospectively, our calculations of potential govern our estimates of TFP.

*fundamental problem is that the use of the filter by non RBC economists imposes a view of the business cycle in which deviations of actual from potential are symmetrically distributed above and below trend.* The same would be true if one ran an OLS regression through the data, and used the line of best fit as an estimate of potential.

This, in my view, is a more serious deficiency than the imposition of a constant between business peaks growth rate of potential reflected in use of the older procedures. From the vantage point of the 1970s it might have appeared to make sense. From a longer run perspective it does not.

Beginning in the late 1970s, most macroeconomics textbooks included a chart showing an upward sloping line on which is superimposed a wavey line. The upward sloping line represents the trajectory of potential. Above the line we have booms, associated with an acceleration of inflation; below the line we have recessions, in which inflation is not accelerating, and during which we experience disinflation or even possibly deflation. Although this retains the idea that trend growth rates are relatively stable at business cycle frequency, the way the wavey line is drawn suggests the view that deviations from trend are symmetrically distributed above and below it.

The responsibility for the current popularity of the use of the HP filter in estimating output gaps cannot be said to lie entirely with Prescott or other RBC proponents, since they had no use for the concept of an output gap as traditionally understood: the economy was always, by definition and by assumption, at potential. He was very clear that HP calculations of trend were just that: a statistical construct. As he said in his Nobel lecture, “trend is not a theoretical concept.”

The problem is that in the hands of scholars who for the most part reject the more untenable aspects of the RBC approach, HP trend calculations have become a stand in for what *is* a theoretical concept: potential output. Let's be clear: Prescott is saying HP trend is just that – it makes no sense to talk about potential as distinguishable from actual, because the two are always equal. But many others, have embraced the HP statistical methodology as a replacement for the old time, judgment ridden trends through peaks methodology, thrilled with the thought that trend growth rate can now be seen as continuously variable, and blind to the violence this approach is doing to what we can conclude historically about business cycles: in particular that deviations are asymmetrically distributed below potential. The appeal of the filter vs trends through peaks approach is understandable. It is handicraft manufacturing vs. industrial production. Dump the data into a computer program, turn the crank, and voila. An estimate of the long run trend in actual may be a reasonable proxy for the unobservable long run trend in potential. It will not necessarily provide a good stand in for annual or quarterly estimates of levels.

Since at least the 1980s, there has been widespread agreement that points above potential are, by definition, associated with periods of accelerating inflation. The trends through peaks approach is premised on this principle, and insures that it is not violated. Let's look at what happens if we allow an HP trend of actual output to morph into our measure of potential.

Robert Gordon (2012) makes available online his annual estimates of real GDP in 2005 dollars for 1875 through 2010. The chart below shows these numbers along with a trend growth line produced using an HP filter and a lambda of 100, commonly recommended for annual data (see Davis, Hanes, and Rhode, 2009).[7] First, it is simply a matter of construction that the sum of

[7] The GDP data have been logged prior to applying the filter.

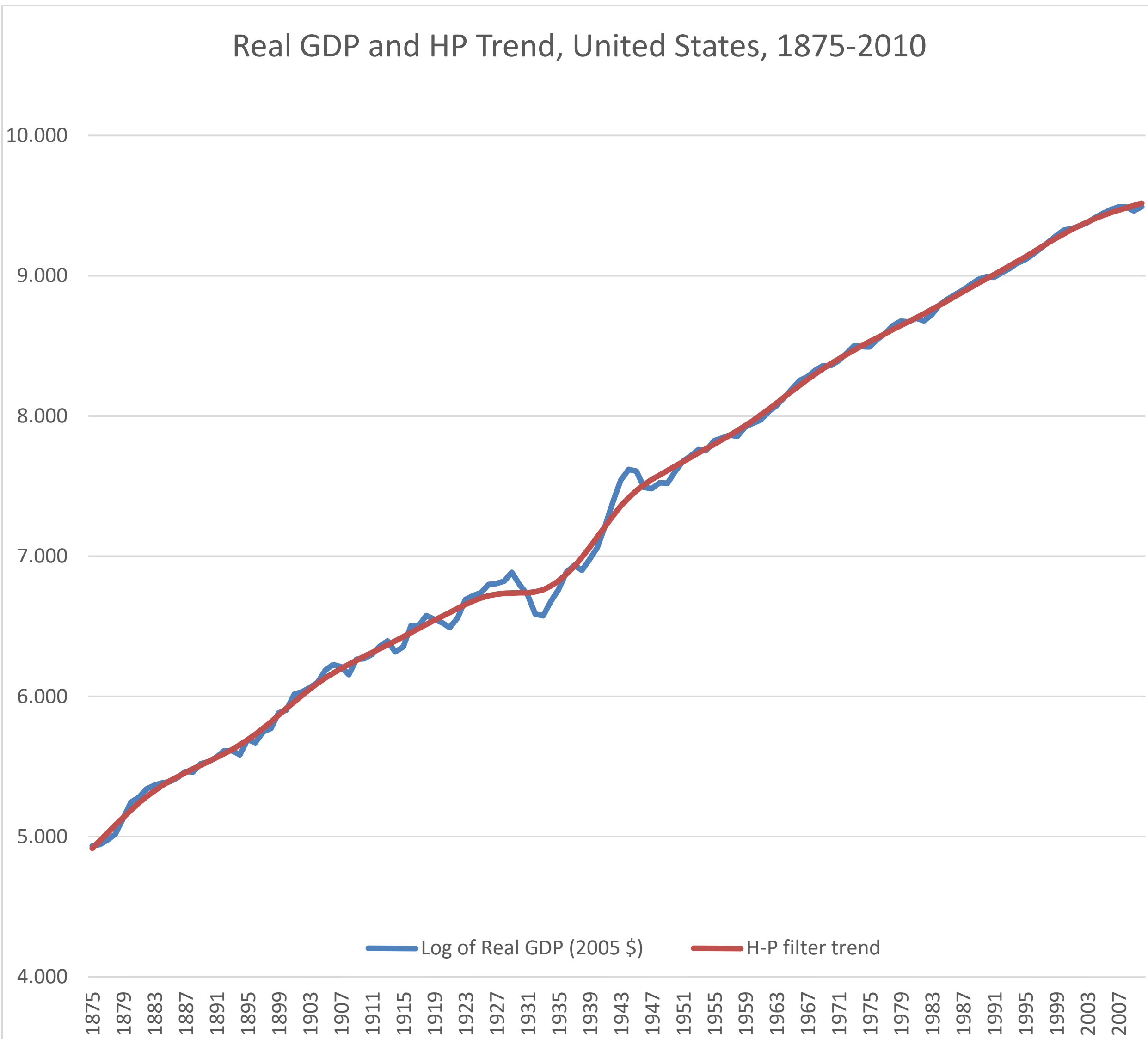
percent deviations above trend (as measured by log differences) will be equal to the sum of deviations below trend. We can then ask, of those years identified as above trend, how many of them were associated with what is by definition a necessary condition for that designation: accelerating inflation? We associate accelerating inflation with above potential because we are saying that in those instances nominal income is rising faster than the real economy can grow, given scientific and technological knowledge, the legal/institutional/regulatory system, the accumulated capital stock, and the characteristics of the labor force including work-leisure preferences.

Here I use Gordon's long series for the implicit GDP deflator, calculating inflation as log differences in levels, and acceleration of inflation as log differences of the log differences. The answer is this. Of the 136 years in the sample, 73 are above trend, 63 below trend. Of the 73 years above trend, 30 are accompanied by *decelerating* inflation.[8]

What does it mean to have a method that conflates a statistical trend calculation with potential output and identifies 30 years experiencing decelerating inflation as above potential? Because the trajectory of potential output is now systematically lower than it would be with a trends through peaks approach, we are likely to have a systematic bias toward tighter monetary and fiscal policy, higher real interest rates, lower employment, and higher unemployment. And because the HP filter will almost invariably show years just prior to a downturn as above trend, we reinforce the tendency to view such years as above potential and therefore unsustainable, even though there is often no evidence of accelerating inflation, and even though the principal

[8] If one is concerned that the quality of data prior to 1929 is too noisy or inaccurate to be used to investigate cyclical phenomena, as Kuznets and Gallman cautioned, we can repeat the exercise, running the HP filter on data from 1929 through 2010. Of those 81 years, 41 are identified as above trend. In 15 of the 41 years, inflation is decelerating.

empirical argument for the position that higher levels of output were unsustainable appears to be simply that they were not sustained. This is not to dispute that an economy can sometimes be above potential – the years of the second world war are examples. But it is to say that the habit of using an HP filter applied to actual data to identify a trend interpreted as the trajectory of potential will misidentify a large number of observations as above potential when they have not in fact satisfied the most basic criterion for being so designated.



This, then, is the most damaging legacy of the RBC initiative. It is not the claim that TFP shocks cause most cyclical variation in the economy. Nor is it the claim that labor supply is very elastic in response to small changes in the real wage rate. These and other such claims have been widely disputed, and to a large degree rejected (witness Mankiw's remarks). The RBC initiative was always about method. Yet perhaps its least anticipated consequence has been the use of a tool championed by its main proponents by those who reject some of the most implausible RBC assumptions and continue to believe that the concept of an output gap is important and useful, yet end up adopting a deformed measure of potential output and thus the output gap.

The proposition that deviations of actual from potential are asymmetrically distributed below potential was central to Friedman's plucking model (1963, 1969). He argued that business cycles phenomena were analogous to what happened when one pulled down on and then released the string of a guitar or violin. The economy would then spring back to potential (what he would later call natural output). His evidence for asymmetry was that there was "no systematic connection" between the size of an expansion and the size of a subsequent contraction, whereas the reverse was true for the relationship between the size of a contraction and the size of a subsequent expansion. He saw no reason to modify his views when he revisited the issue three decades later (1993, pp. 171-72).[9]

Asymmetry is also cautiously endorsed by deLong and Summers, who suggest

> ….some empirical support for viewing business cycles as gaps rather than as cycles around supply driven trends. The existence of cyclical asymmetries, the correlations of constructed gaps with observed unemployment, and the stronger response of output to negative than to positive monetary shocks together suggest that the gaps view may provide

[9] Friedman also argued for another empirical regularity: the further down you pulled the string (the deeper the recession), the faster it would spring back (the more rapid the recovery). In propounding this regularity, he did not fully consider the possible confounding effects of financial crises and balance sheets, to which we have now become more attuned.

> a more accurate characterization of fluctuations than does the more standard view of fluctuations as near symmetric cycles around unique equilibrium trend levels of output and unemployment…. Asymmetry fits more naturally into a framework that sees fluctuations as lapses beneath potential than into one that sees them as cycles around trend (deLong and Summers, 1988 pp. 436, 446).

The "more standard view" is shared by RBC proponents and many of their New Keynesian "opponents." It is in part the legacy of the widespread penetration of a methodological tool – the HP filter, championed by RBC proponents – throughout macroeconomics. RBC proponents like Prescott are clear that the trend series produced by the filter does not correspond to a theoretical concept. Yet the filter has been repurposed by many in the profession to provide estimates of a series that does.

**Conclusion**

In Great Depressions of the Twentieth Century, Timothy J. Kehoe and Prescott argue that "Collectively, (the papers in the volume) indicate that government policies that affect productivity and hours per working-age person are the crucial determinants of the great depressions of the twentieth century." One reason that real business cycle models and research that exploits them have apparently made so little impact in the pages of economic history journals is simply that many macroeconomists and most macroeconomic historians find such claims to be, for lack of a better word, preposterous. Perhaps that is too strong a word, although Summers used language such as "defies credibility" and "simply absurd" to describe attempts to apply RBC models to the Great Depression (Summers, 1986, p. 26).

RBC proponents persuaded themselves, and a number of others, that they had developed a new scientific methodology. But it was never clear whether it was about hypothesis testing, or estimation, or something else entirely. Much faculty and student attention is now devoted to explaining the outline of the RBC research program as well as critiques of it. In the process, the balance of macroeconomics research shifts subtly from studies of the characteristics of the economy to studies of models per se. RBC methods tells us nothing about whether the principal source of business cycle fluctuations are aggregate demand or aggregate supply shocks. RBC models contributed nothing to the diagnosis and treatment of the financial crisis of 2007-2008. RBC models have added little or nothing to our understanding of the Great Depression.

But the RBC initiative has done more than simply diverting intellectual effort away from activities that might have been more productive. If the initiative will probably not succeed in shunting the car of economic science onto the wrong track for a hundred years, it has succeeded in damaging how we understand and measure key theoretical concepts: potential output and the output gap. Relatively few people today believe that negative supply shocks were responsible for the Great Depression. Relatively few believe that that labor supply is subject to a high intertemporal elasticity of substitution. If Kydland and Prescott wanted the rest of the profession to accept these propositions they have largely failed.

But method was always front and center in the RBC initiative. As Prescott said in his Nobel lecture, “What I am going to describe for you is a revolution in macroeconomics, a transformation in methodology that has reshaped how we conduct our science (2006, p. 205). … In fact, the meaning of the word macroeconomics has changed to refer to the tools being used rather than just to the study of business cycle fluctuations (2006, p. 204).” The evaluative methodology proposed – casual measures of the degree of correspondence between standard deviations/comovements in model output and in actual macroeconomic aggregates, has no bearing on model validity. RBC proponents both claim that it does and deny that it should.

Above and beyond this, the championing of the HP filter and the diffusion of its use among empirical economists and economic historians has proved damaging in a subversive way. Many economists have adopted the HP filter as a convenient, modern, sophisticated means of providing an estimate of the past trajectory of potential. A statistical tool advanced to separate trend from cycle, accompanied with a clear statement by Prescott that trend did not correspond to any theoretical concept, has been repurposed by many macroeconomists who, unlike Prescott, do not believe we are always by definition at potential, and still believe the concept of an output gap is useful and important in macroeconomics.

This process has subtly weakened the insistence that periods above potential be accompanied by accelerating inflation, has lowered estimates of potential from where they would be if that criterion were religiously adhered to. Inflation, and particularly variability in the inflation rate, can be quite damaging to an economy, and in extreme form will shred the social and political fabric of a country. The story of Weimar Germany and the consequences of runaway inflation has been often told. Potential output is a real limit in any actual economy. We will know when we get there if and when we see an acceleration of inflation. But corrupted measures of potential encourage us to stop before we get close, providing justification for restrictive trajectories of monetary and fiscal policy.

SONY withdrew AIBO from production in 2006, following sales of 150,000 units, and stopped providing all maintenance services in 2014 (Mochizuki and Pfanner, 2015). Something roughly analogous has happened with the original RBC models. New Keynesian models, driven by different priors, are still in production. There are many similarities between the robot canine and these classes of macro models. We can presume that SONY engineers took pride in their creations, and considerable evidence that those who played with them got much enjoyment, and, indeed, found them aesthetically pleasing. One can say something similar about the creators and users of RBC models and methods.

There is however an important contrast. None of the creators and users of AIBO believed that their success in creating a machine that did a passable job mimicking the behavior of an actual mutt had any bearing on a possible claim that the servo motors and electromechanical systems were a good model of a dog's innards. In contrast, creators and users of RBC models believe that their success in mimicking certain features of macro time series provided evidence bearing on the validity of the maintained hypotheses that supply shocks cause most business cycles, that markets clear, that the economy in the aggregate experiences high intertemporal substitutability of leisure, and that all unemployment is voluntary.

My own view is that New Keynesian priors – in particular rejection of the view that all cycles are caused by technology shocks, are more realistic. But the exercises in New Keynesian model building have neither decreased nor increased my confidence in these priors, which are based on other evidence. It is hard to see how they could. In developing a broader range of DSGE models, the NK methods are at heart RBC. As Nakamura's work makes particularly clear, RBC standards for evaluation do not allow models to be tested in any meaningful way against alternate views about what might cause business cycles. They do not provide a space for the kind of results that might change priors. The enterprise, as was true of the original RBC models, is largely about appearing to confirm prior beliefs. Ultimately the NK models will suffer the same neglect as do the original RBC versions.

A fundamental priority of macroeconomic historians has always been to understand and explain the data, and to use this analysis to inform historical narrative. We have used theory where helpful, and, where applicable, viewed history as a means of disciplining and refining theory. The decade of the 1970s posed challenges to then prevailing neo-Keynesian orthodoxy. Aggregate supply always lay in the background, but was often abstracted from for short term analysis. For some episodes (1974-75 in particular) this was unwarranted; a greater attention to aggregate supply made sense. The formalization of an expectations augmented Philips curve provided a coherent account of stagflation within a framework that had been usefully applied to many earlier macroeconomic episodes.

The RBC initiative, in contrast, with assumptions aimed at setting the theory clock back to the 1920s, and a methodology imposing aesthetic criteria and often casual standards for model evaluation, exploited the opening provided by the 1970s to advance a radical new way of doing macroeconomics. In maintaining priorities, economic historians have been right, by and large, to keep some distance. We should continue to be aware of the unusual evaluative methods employed and the damage the diffusion of RBC empirical tools has done to our understanding of the macroeconomy – damage for which many who reject the more radical claims of RBC proponents also bear some responsibility.

BIBLIOGRAPHY


Bean, Louis H. 1945. "Postwar Output in the United States at Full Employment." Review of Economics and Statistics 27 (November): 202-203.

Bordo, Michael D., Christopher J. Erceg and Charles L. Evans. 2000 "Money, Sticky Wages, and the Great Depression." The American Economic Review 90 (December): 1447-1463.

Calomiris, C. W. and Hanes, Christopher. 1995. "Historical Macroeconomics and Macroeconomic History." In Kevin D. Hoover, ed., Macroeconometrics: Developments, Tensions, and Prospects. Dordercht: Kluwer, pp. 351-416.

Cole H. L. and L. E. Ohanian. 1999. "The Great Depression in the United States from a Neoclassical Perspective". Federal Reserve Bank of Minneapolis Quarterly Review 23: 2-24.

Cole H. L. and L. E. Ohanian. 2000. "Re-Examining the Contributions of Money and Banking Shocks to the U.S. Great Depression", Research Department Staff Report 270, Federal Reserve Bank of Minneapolis.

Cole H. L. and L. E. Ohanian. 2002. "The US and UK Great Depressions through the Lens of Neoclassical Growth Theory", American Economic Review 92: 28-32.

Cole H. L. and L. E. Ohanian. 2004. "New Deal Policies and the Persistence of the Great Depression: A General Equilibrium Analysis." Journal of Political Economy 112: 779-816.

deLong, J. Bradford and Lawrence H. Summers. 1988. "How Does Macroeconomic Policy Affect Output?" Brookings Papers on Economic Activity 2: 433-494.

Davis, Joseph H., Christopher Hanes, and Paul W. Rhode. 2009. "Harvests and Business Cycles in Nineteenth-Century America." Quarterly Journal of Economics (November): 1675-1727.

Despres, Emile, Albert G. Hart, Milton Friedman, Paul A. Samuelson and Donald H. Wallace. 1950. "The Problem of Economic Instability." The American Economic Review 40 (September):501-538.

De Vroey, Michel R. and Luca Pensieroso. 2006. "Real Business Cycle Theory and the Great Depression: The Abandonment of the Abstentionist Viewpoint." Contributions to Macroeconomics 6 Article 13.

Field, Alexander J. 1985. "On the Unimportance of Machinery." Explorations in Economic History 22 (October): 378-401.

Field, Alexander J. 1992. "Uncontrolled Land Development and the Duration of the Depression in the United States," Journal of Economic History 52 (December): 785-805.

Field, Alexander J. 2010. "The Procyclical Behavior of Total Factor Productivity, 1890-2004." Journal of Economic History 70 (June): 326-50.

Field, Alexander J. 2011. A Great Leap Forward: 1930s Depression and US Economic Growth. New Haven: Yale University Press.

Field, Alexander J. 2015. "The Saving and Loan Insolvencies in the Shadow of 2007-09." Working Paper.

Field, Alexander J. 2016. "The Taylor Rule in the 1920s." Working Paper.

Friedman, Milton. 1953. "The Methodology of Positive Economics." In Essays in Positive Economics. Chicago: University of Chicago Press.

Friedman, Milton. 1997. "Computational Experiments." Journal of Economic Perspectives 11 : 209-10.

Friedman, Milton. 1969. The Optimum Quantity of Money and Other Essays. Chicago: Aldine.

Friedman, Milton. 1993. "The "Plucking Model" of Business Fluctuations Revisited." Economic Inquiry 31 (April): 171-77.

Galí, Jordi, J. David López-Salido, and Javier Vallés. 2007. "Understanding the Effects of Government Spending on Consumption." Journal of the European Economic Association 5 (March): 227-270.

Gordon, Robert J. 2012. "Time Series Data for the U.S. Economy: 1875-2010." Available at http://www.pearsonhighered.com/gordon. Accessed December 15, 2015.

Hagen, Everett E. 1945. "Postwar Output in the United States at Full Employment." Review of Economics and Statistics 27 (May): 45-59.

Hansen, Lars Peter and James Heckman. 1996. "The Empirical Foundations of Calibration." Journal of Economic Perspectives 10 (Winter): 87-104.

Hall, Robert E. 1999. "Labor Market Frictions and Employment Fluctuations." In In Michael Woodford and John Taylor, eds. Handbook of Macroeconomics, v. 1b. Ch. 17 Amsterdam: Elsevier, pp. 1138-1170

Hartley, James E. 1997. The Representative Agent in Economics. London: Routledge.

Hartley, James E., Kevin D. Hoover and Kevin D. Salyer. 1997. "The Limits of Business Cycle Research: Assessing the Real Business Cycle Model. Oxford Review of Economic Policy 13: 34-54.

Hoover, Kevin D. 1995. "Facts and Artifacts: Calibration and the Empirical Assessment of Real Business Cycle Models." Oxford Economic Papers 47: 24-44.

Kehoe, Timothy and Edward Prescott, eds. 2007. Great Depressions of the Twentieth Century. Federal Reserve Bank of Minneapolis.

Kimball, Miles and Mathew Shapiro. 2008. "Labor Supply: Are the Income and Substitution Effects Both Large or Small?" NBER Working Paper 14208 (July).

King, Robert G. and Sergio T. Rebelo. 1999. "Resuscitating Real Business Cycles." In Michael Woodford and John Taylor, eds. Handbook of Macroeconomics, v. 1. Amsterdam: Elsevier, pp. 928-1007.

Kydland, Finn. 2004. "Quantitative Aggregate Theory." Nobel Prize Lecture. Available at http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2004/kydland-lecture.html.

Kydland, Finn and Edward Prescott. 1977. “Rules rather than Discretion: The inconsistency of Optimal Plans”. Journal of Political Economy 85: 473-490.

Kydland, Finn and Edward Prescott. 1982. “Time to build and aggregate fluctuations.” Econometrica 50: 1345-1371.

Kydland, Finn and Edward Prescott. 1991. “The Econometrics of the General Equilibrium Approach to Business Cycles.” Scandinavian Journal of Economics 93: 161-78.

Long, John B. J. and Charles I. Plosser. 1983. “Real Business Cycles.” Journal of Political Economy 91: 39-69.

Mankiw, N Gregory. 1989. "Real Business Cycles: A New Keynesian Perspective." Journal of Economic Perspectives 3: 79-90.

Mochizuki, Takashi and Eric Pfanner. 2015. “In Japan, Dog Owners Feel Abandoned as Sony Stops Supporting ‘Aibo’.” Wall Street Journal (February 11):

Musgrave, Richard R. 1945. ‘Alternative Budget Policies for Full Employment.” The American Economic Review 35 (June): 387-400.

Nakamura, Emi. 2009. “Deconstructing the Success of Real Business Cycle Models.” Economic Inquiry 47 (October): 739-753.

Parker, Randall E. 2007. The Economics of the Great Depression: A Twenty First Century Look Back at the Economics of the Interwar Era. Cheltenham and Northampton, MA.: Edward Elgar.

Plosser, Charles. 1989. “Understanding Real Business Cycles.” Journal of Economic Perspectives 3 (Summer): 51-77.

Prescott, Edward C. 1986. “Theory Ahead of Business Cycle Measurement.” Federal Reserve Bank of Minneapolis Quarterly Review (Fall): 9-22.

Prescott, Edward C. 2006. “Nobel Lecture: The Transformation of Macroeconomic Policy and Research: Journal of Political Economy 114 (April): 203-235.

Prescott, Edward C. 2016. “RBC Methodology and the Development of Aggregate Economic Theory.” Minneapolis Federal Reserve Bank Staff Report 527.

Thurow, Lester C and Lance Taylor. 1966 The Review of Economics and Statistics 48, No. 4 (November): 351-360

Royal Swedish Academy of Science. 2004. “Finn Kydland and Edward Prescott’s Contribution to Dynamic Macroeconomics: The Time Consistency of Economic Policy and the Driving Forces Behind Business Cycles.” Available at *http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2004/advanced-economicsciences2004.pdf*.

Schelling, Thomas. 1947. “Capital Growth and Equilibrium. ”The American Economic Review 37 (December): 864-876.

Stadler, George W. 1994. “Real Business Cycles.” Journal of Economic Literature 32 (December): 1750-1783.

Summers, Lawrence . 1986. “Some Skeptical Observations on Real Business Cycle Theory.” Federal Reserve Bank of Minneapolis Quarterly Review (Fall): 23-27.

Watson, Mark. 1993. “Measures of Fit for Calibrated Models.” Journal of Political Economy 101 (December): 1011-1041.

Wren-Lewis, Simon. 2016. “Unravelling the New Classical Counterrevolution.” Review of Keynesian Economics 4: 20-35.