Federal Reserve Bank of Minneapolis
Research Department Staff Report ___

April 2008

New Keynesian Models Are Not Yet Useful for Policy Analysis*

V. V. Chari

Federal Reserve Bank of Minneapolis
and University of Minnesota

Ellen R. McGrattan

Federal Reserve Bank of Minneapolis
and University of Minnesota

Patrick J. Kehoe

Federal Reserve Bank of Minneapolis
and University of Minnesota

ABSTRACT

In the 1970s macroeconomists often disagreed bitterly. Macroeconomists have now largely converged on method, model design, and macroeconomic policy advice. The disagreements that remain all stem from the practical implementation of the methodology. Some macroeconomists think that New Keynesian models are on the verge of being useful for quarter-to-quarter quantitative policy advice. We do not. We argue that the shocks in these models are dubiously structural and show that many of the features of the model as well as the implications due to these features are inconsistent with microeconomic evidence. These arguments lead us to conclude that New Keynesian models are not yet useful for policy analysis.

*The views expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

Viewed from a distance, academic macroeconomists are like Tolstoy's happy families: We are all alike, at least in the sense that we use the same methodology, work with similar models, and often even prescribe similar policies. Viewed up close, however, we are more like Tolstoy's unhappy families: We are all unhappy with the state-of-the-art models of monetary policy, just each in our own way. The common thread among the disparate forms of unhappiness is that we are frustrated with our inability to give sound quarter-by-quarter quantitative macroeconomic policy advice. The most important practical difference among us is that some of us are willing to admit to this inability and others are not.

Modern macroeconomists use essentially one methodology. We all agree that to do serious policy analysis, we need a model in which the main elements are arguably invariant to the class of policy interventions being considered.

Macroeconomic models are also similar. In practice, most macroeconomists now analyze policy using dynamic stochastic general equilibrium (DSGE) models. These models can be so generally defined that they incorporate all types of frictions, including various ways of learning, incomplete markets, imperfections in markets, spatial frictions, and so on. The only practical restriction from these models is that they specify an agreed-upon language by which we communicate. A standard aphorism is that if you have a coherent story to propose, then you can do so in a suitably elaborate DSGE model.

Modern macroeconomists also broadly concur on the desirable properties of monetary policy. First, the success of policy depends on policymakers' commitment. Second, interest rates and inflation rates should be kept low on average. More practically, most macroeconomists are comfortable with some form of inflation targets which have some well-defined escape clauses. The concurrence on policy is not surprising since we are all using DSGE models.

This agreement on method, model design, and macroeconomic policy advice is remarkably different from the discord of the 1970s and 1980s. The *Old Keynesian* view is eloquently and forcefully summarized by Modigliani (1977, p. 1), who argues that the fundamental practical policy implication that Old Keynesians agree on is that the private economy "*needs* to be stabilized, *can* be stabilized, and therefore *should* be stabilized by appropriate monetary and fiscal policies." From this perspective, a commitment to low average inflation rates is an unnecessary and harmful constraint on wise policy. Neoclassical economists, like Robert Lucas and Edward Prescott, agree. They hold that a commitment to rules is essential for good economic performance and that good rules are those that ensure low and stable inflation.

How did this convergence happen? Three considerations drive us to the view that the main convergence was from the Old Keynesian view to the *New Keynesian* view, which turns out to be very close to where the Neoclassical view has been all along. First, since modern macroeconomists use equilibrium models with forward-looking private agents, a commitment to rules is essential for good economic performance. Second, even in the frictionless version of all modern models, efficient allocations fluctuate sizably. In this sense, even under optimal policy, the model will display sizable business cycle fluctuations, and eliminating all of these fluctuations is bad policy. Third, New Keynesian models typically incorporate sticky prices or wages, and optimal monetary policy in such models typically keeps inflation low and stable in order to avoid sectoral misallocations.

But, of course, we don't agree on everything. The disagreements all stem from the practical implementation of the methodology. So far, for example, there is little consensus on what primitive interpretable shocks account for the bulk of business cycle fluctuations. There is relatively more consensus on the necessary reduced-form shocks, which we term *wedges* in earlier work (Chari, Kehoe, and McGrattan (2007), henceforth *CKM*). At face value, these wedges look like time-varying productivity, labor income taxes, investment taxes, and government consumption. We thus label the wedges *efficiency wedges*, *labor wedges*, *investment wedges*, and *government consumption wedges.*

The idea of our *business cycle accounting* approach is that studying these wedges in a prototype model can help guide researchers about where to introduce frictions into their detailed models. CKM show how to guide theory by using an *equivalence result* that a large class of detailed models, including models with various types of frictions, are equivalent to a prototype model with various types of time-varying wedges which distort the equilibrium decisions of agents operating in otherwise competitive markets. The class of detailed models that give rise to patterns of wedges similar to those in the data are deemed the most promising ones.

From a study of both the Great Depression and postwar business cycles, CKM determine that the efficiency and labor wedges account for the vast bulk of movements in the macro aggregates. Using the equivalence result, we suggest classes of promising detailed models which give rise to these wedges. One class of promising models are New Keynesian models with sticky prices and sticky wages because these models have the potential to generate large labor wedges from primitive interpretable shocks.

Unfortunately, as we will show here, the state-of-the-art New Keynesian model, represented by the Smets and Wouters (2007) model, fails to live up to its promise. The reason it fails is not that it cannot produce efficiency or labor wedges, but that it does so by adding on such dubious features that it makes little sense to think of the model as structural. Hence, the model does not provide any more insight into the source of the wedges than does the original prototype model of CKM. Moreover, since the Smets-Wouters model is not structural, it cannot be used for policy.

Smets and Wouters (2007) and others in the New Keynesian camp may well disagree with our assessment of their workhorse model. That disagreement stems from a deeper disagreement about model building (that does not split neatly along traditional New Keynesian–Neoclassical lines). There are two basic traditions in model building and assessment. One is to keep the model very simple, keep the number of parameters small and well-motivated by micro facts, and put up with the reality that such a model neither can nor should fit most aspects of the data. Such a model can still be very useful in clarifying how to think about policy. Typical examples are the general equilibrium models of optimal fiscal policy of Lucas and Stokey (1983) which make clear general principles, such as the optimality of smoothing distortions over time and across states. When these models are quantitatively implemented, a simple rule of thumb used to discourage the adding of free parameters is that every time a new parameter is added, some new micro evidence to discipline that parameter should be added as well.

The other tradition, typified by the work of Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007), emphasizes the need to fit macro aggregates well. The urge to fit these aggregates well leads researchers in this tradition to add many more features and shocks and then try to use the same old aggregate data as before to estimate the associated new parameters without the discipline of microeconomic evidence.

We label many of these added parameters *structurally dubious* for two reasons. First, because the addition of new parameters is undisciplined, the added features and shocks are often hard to view as structural with respect to policy interventions. Second, the added features and shocks are typically of such a reduced-form nature that they are consistent with several structural interpretations. The policies associated with each structural interpretation are often very different. In this sense, even though the models in this tradition, by construction, fit the macro aggregates well, they are not yet useful for policy analysis.

Almost every feature and shock added to the Smets-Wouters version of the New Keynesian model, beyond that of the elegant version by Yun (1996), is structurally dubious. Here we focus on only the most egregious of them. The wage-markup shock accounts for the bulk of the fluctuations. As we demonstrate, putting that shock into the model is equivalent to mechanically sticking in a labor wedge. It is equally as interpretable (technically unidentified) as fluctuations in the bargaining power of unions or fluctuations in the value of leisure of consumers. We show that policy implications vary drastically depending on what interpretation is adopted.

Furthermore, either interpretation seems strained. In the bargaining power view, a contagious attack of greediness among workers leads them to demand higher wages. In general equilibrium, this attempt is frustrated, and these workers simply bid themselves out of jobs. In the fluctuating value of leisure view, a contagious attack of laziness among workers leads them all to take vacations by quitting, thus causing the economic downturn. Many macroeconomists will find both interpretations uninteresting and hence will find the model not an attractive guide for policy. For those who find these interpretations promising, the task is to find micro evidence to convince the profession that these shocks are both invariant to policy and interpretable enough so that we know whether policy should try to offset or accommodate them.

We then turn to two other structurally dubious features: backward indexation and the common specification of the Taylor rule. We argue that they are both inconsistent with the data. Consider the backward indexation of prices. This feature is a mechanical way for the model to match the persistence of inflation. We show that this feature is flatly inconsistent with the micro data on prices. Consider next the Taylor rule, which is a specification of how the Federal Reserve sets the short-term nominal rate as a function of what it observes. We argue that the Smets-Wouters specification, which follows a long tradition in assuming the short rate is stationary and ergodic, is incapable of generating anything close to the observed behavior of the long-term nominal rate. Since the behavior of the long-term rate reflects in an important way how the policy instrument, the short rate, affects the real side of the economy, misspecifying this relationship leads to a very inaccurate assessment of policy.

We argue that the last two problems, the backward indexation and the dubious specification of Fed policy, may be linked. Once we specify the Fed's policy as having a random walk–like component, the resulting model can fit the aggregates without the structurally dubious backward indexation. In particular, the persistence of inflation seen in the data naturally follows from the persistence of policy, instead of having to be tacked on to the model in a mechanical way. To see why getting the true structure correct is critical for policy, consider the costs of a disinflation. With backwardly indexed prices, these costs are huge; without them, the costs are tiny. Hence, here is another example of how tacking on mechanical, structurally dubious features can improve a model's fit but render it not yet useful for policy analysis.

We then briefly examine why two other shocks, the risk premium shock and the exogenous spending shock, are also structurally dubious.

We have focused our attention on the Smets and Wouters (2007) model because it is the most serious and ambitious attempt at translating New Keynesian ideas into a coherent fully-specified general equilibrium model. We have high regard for this attempt because the authors were willing to venture out from the narrow issue of whether a model can fit a few impulse responses to the broader issue of building a model that can generate the business cycle. In this sense our lengthy and detailed criticism of this work shows our high regard for this attempt.

Our overall message is that there are severe problems with this general approach. If the structurally dubious features, such as the so-called markup shocks or the backward indexation, are added more for mechanical reasons, to help fit, rather than for sound economic reasons, then the resulting policy implications may be rendered useless.

1. Consensus on Wedges, But Not Primitive Shocks

In order to do policy analysis, we need a structural model. Specifically, we need the elements of the model—including the shocks—to be reasonably understood to be invariant with respect to the policy interventions considered. Beyond that, though, we also need the shocks to be interpretable, so that we know whether they are what could be thought of as "good shocks" that policy should accommodate or "bad shocks" that policy should offset.

To date, there is little consensus on what primitive interpretable shocks account for the bulk of business cycle fluctuations. There is more consensus on reduced-form shocks, termed *wedges* by CKM. The idea of CKM is that when researchers build detailed, quantitative models of economic fluctuations, they face hard choices about where to introduce frictions into their models in order to allow the models to generate business cycle fluctuations similar to those in the data.

Our method of providing guidance to such researchers, referred to as *business cycle accounting*, has two components: an equivalence result and an accounting procedure. The *equivalence result* is that a large class of models, including models with various types of frictions, are equivalent to a prototype model with various types of time-varying wedges that distort the equilibrium decisions of agents operating in otherwise competitive markets. At face value, these wedges look like time-varying productivity, labor income taxes, investment taxes, and government consumption. We thus label the wedges *efficiency wedges*, *labor wedges*, *investment wedges*, and *government consumption wedges*.

The *accounting procedure* also has two components. It begins by measuring the wedges, using data together with the equilibrium conditions of a prototype model. The measured wedge values are then fed back into the prototype model, one at a time and in combinations, in order to assess how much of the observed movements of output, labor, and investment can be attributed to each wedge, separately and in combinations. The motivation for calling our method *business cycle accounting* is that, by construction, all four wedges account for all of these observed movements.

We apply this procedure to two U.S. business cycle episodes: the most extreme in U.S. history, the Great Depression (1929–39), and a postwar recession, the 1982 recession, which can be thought of as representative of postwar recessions. For both the Great Depression period and the 1982 recession, we find that the efficiency wedge and the labor wedge accounted for the vast bulk of business cycle movements. The investment wedge played no useful role in the Great Depression and a limited role in the postwar recession. The government spending wedge played essentially no role in either episode.

A. An Introduction to Business Cycle Accounting

To get the gist of this procedure, consider a prototype economy, which is a standard one-sector growth model with four exogenous stochastic processes: the *efficiency wedge* A_t, the *labor wedge* $1 - \tau_{lt}$, the *investment wedge* $1/(1 + \tau_{xt})$, and the *government consumption wedge* g_t. (For simplicity in this exposition of the procedure, we abstract from population growth and labor-augmenting technical change, but in the quantitative work, we do not. See CKM for details.)

In this economy, consumers maximize expected utility over per capita consumption c_t and per capita labor l_t,

$$E_0 \sum_{t=0}^{\infty} \beta^t U\big(c_t, 1 - l_t\big),$$

subject to the budget constraint

$$c_t + (1 + \tau_{xt})\, x_t = (1 - \tau_{lt})\, w_t l_t + r_t k_t + T_t$$

and the capital accumulation law

$$k_{t+1} = (1 - \delta) k_t + x_t, \tag{1}$$

where k_t denotes the per capita capital stock, x_t per capita investment, w_t the wage rate, r_t the rental rate on capital, β the discount factor, δ the depreciation rate of capital, and T_t per capita lump-sum transfers. Notice that in this prototype economy, the efficiency wedge resembles a blueprint technology parameter, and the labor wedge and the investment wedge resemble tax rates on labor income and investment.

The equilibrium of this prototype economy is summarized by the resource constraint,

$$c_t + x_t + g_t = y_t, \tag{2}$$

where y_t denotes per capita output, together with

$$y_t = A_t F\big(k_t, l_t\big), \tag{3}$$

(4) $$\frac{U_{lt}}{U_{ct}} = (1 - \tau_{lt})\, A_t F_{lt}, \text{ and}$$

(5) $$U_{ct}\,(1 + \tau_{xt}) = E_t\left[\beta U_{ct+1}\{A_{t+1} F_{kt+1} + (1 - \delta)(1 + \tau_{xt+1})\}\right],$$

where, here and throughout, notations like U_{ct}, U_{lt}, F_{lt}, and F_{kt} denote the derivatives of the utility function and the production function with respect to their arguments.

To execute this accounting procedure we need functional forms for preferences and technology. We use forms popular in the business cycle literature $U(c, l) = \log c + \psi \log(1 - l)$ and $F(k, l) = k^{\alpha} l^{1-\alpha}$. With these forms, the efficiency and labor wedges are given by the simple static relationships

$$A_t = \frac{y_t}{k_t^{\alpha} l_t^{1-\alpha}} \text{ and } 1 - \tau_{lt} = \left(\frac{\psi}{1 - \alpha}\right)\left(\frac{c_t}{y_t}\right)\left(\frac{l_t}{1 - l_t}\right).$$

Hence, the realizations of these wedges can be obtained once we specify the parameters α and ψ. The government consumption wedge is defined residually from (2) and can be obtained once we take a stand on our measures of consumption, investment, and output. Note that since the actual U.S. economy is an open economy with nontrivial net exports, the residually defined government consumption wedge will correspond to the sum of government spending and net exports.

The subtle wedge to measure is the investment wedge. This wedge shows up in the intertemporal Euler equation. Hence, we need to take a stand on conditional expectations in order to measure this wedge. Briefly, to do so, we specify a vector autoregressive process on the four wedges and estimate the parameters of this process using maximum likelihood. (See CKM for details.)

B. The Key Wedges: Efficiency and Labor

To get some feel for why macroeconomists tend to agree that efficiency and labor wedges are the key to reduced-form shocks, consider what our accounting procedure implies for the longest recession in U.S. history: the Great Depression.

In Figure 1, we report on U.S. output (relative to trend) and the measured efficiency and labor wedges for the Great Depression period from 1929 to 1939. We see that the underlying distortions that manifest themselves as efficiency and labor wedges substantially worsened from 1929 to 1933. After 1933, the efficiency wedge recovers, so that by 1939, it is back to trend; but the labor wedge does not recover and is well below trend in 1939.

In Figure 2, we plot the 1929–39 data for U.S. output, labor, and investment along with the model's predictions for those variables when the model includes just one wedge. In this figure, we see that with the efficiency wedge alone, the model captures most of the decline in output and investment and only a small fraction of the movement in labor. We also see that with the labor wedge alone, the model captures almost all of the movements in labor but only a modest fraction of the movements in output and investment. As we show in CKM, the model with both the efficiency and labor wedges capture almost all of the movements in these three series, and the investment and government consumption wedges do very little.

At the risk of somewhat oversimplifying issues, we think these results can be understood in a very intuitive manner. The idea of the estimation procedure is that it must choose the magnitude of the efficiency, labor, and investment wedges so that the model generates data similar to those observed for the output, labor, and investment series. (From now on, we ignore the government consumption wedge since it does essentially nothing.) The procedure's only choices are what time-varying linear combinations of these wedges it must choose. We start by asking what effects each wedge by itself has on these three series.

The efficiency wedge, which in the model works identically to a productivity shock, works as follows. Since it is highly serially correlated, a drop in this wedge signals that productivity is low now and, on average, will remain low in the near future. The optimal response is to have a large drop in investment. Output naturally drops, mainly because the same level of inputs make less output when productivity drops. This shock, however, has only minor effects on the labor input. Intuitively, we know that since this wedge is very persistent, there is little incentive to intertemporally substitute labor—the marginal product of labor is low now, but since it is expected to remain low, there is little tilt away from the present and toward the future.

The labor wedge, which in the model works like a tax on wage income, works as follows. A drop in this wedge, which is akin to a rise in the tax on labor, directly discourages work through the standard static consumption-leisure trade-off: all else equal, one unit of labor devoted to work produces less consumption than before, so that it is optimal to work less for a given consumption level. Because a drop in this wedge leads to a lowering of the labor input, it also leads to a modest fall in output. (In the production function, l is raised to the power $1-\alpha$, which is less than 1; so the direct effect of a 1 percent fall in l is only a $1-\alpha$ percent fall in y.)

The investment wedge, which is akin to a tax on investment, works as follows. A worsening of this wedge, namely, a lowering of $1/(1+\tau_x)$, corresponds to an increase in the tax on investment. Since neither the labor input nor the capital input into production is much affected by this wedge, output produced, given by the right side of the resource constraint (2), is not much affected. The main effect of an increase in this wedge is that resources devoted to investment are lowered, and those devoted to consumption are increased. Mechanically, if investment x_t in (2) drops, and neither y_t nor g_t move much, then what must happen is that c_t rises to offset this fall. Thus, the investment wedge causes consumption and investment to move in opposite directions.

Now the estimation procedure is given data in which output and investment fall together, with investment falling much more (percentage-wise) than output. If the model tried to blame most of the fall in investment on the investment wedge, then it would end up predicting a large boom in consumption in, say, the middle of the Great Depression, which, given the resource constraint (2), and the data presented, clearly did not happen. Hence, the investment wedge is out as a major source of the downturn, and the estimation procedure is left trying to figure out what weights to put on the efficiency and labor wedges in order to justify the data it is given.

Suppose the estimation procedure tries to blame all of the movements in these series on the labor wedge. In particular, suppose that the procedure chooses the labor wedge to reproduce the series on labor exactly. Then the problem will be that the model predicts falls in output and investment much smaller than are seen in the data. Likewise, suppose the procedure tries to blame all of the movements in these series on the efficiency wedge. In particular, suppose the procedure chooses the efficiency wedge so as to reproduce the series on output exactly. The model will then vastly underpredict the fall in labor.

Hence, what the procedure does is to choose to blame most of the fall in output on the efficiency wedge, and the resulting wedge will then account for most of the fall in investment.

To account for the fall in labor, the procedure is forced to choose a fairly large labor wedge. This labor wedge then also leads output and investment to fall enough so that combined the efficiency and labor wedges account for almost all of the movements in the three series.

2. Generating the Observed Pattern of Wedges from Detailed Models

We turn now to discussing how a researcher, armed with the knowledge that the efficiency and labor wedges account for most of the business cycle fluctuations, can use equivalence results between a prototype economy and detailed economies to isolate promising classes of models of the business cycle.

A. The Efficiency Wedge: Frictions, Not Mismeasurement

Consider first the efficiency wedge. The massive drop in the efficiency wedge of nearly 20% relative to trend observed in the 1929–33 period make it *a priori* difficult to interpret as literally a change in the blueprint technologies. CKM argue that a more promising interpretation is that underlying frictions either within or across firms cause factor inputs to be used inefficiently. These frictions often show up as aggregate productivity shocks in a prototype economy similar to ours. We also argue that it is extremely doubtful that this wedge can be written off as simple mismeasurement.

Frictions

To be concrete, CKM consider a simple detailed economy with input-financing frictions and show how variations in these frictions show up in the prototype economy as efficiency wedges. The basic idea is that firms must borrow to pay for an intermediate input in advance of production. One type of firm, call them *small firms*, are more financially constrained than another type, call them *large firms*, in that the small firms must pay a higher rate on borrowing than the large firms. CKM show that fluctuations in the financial constraints across these two types of firms show up in the prototype economy as efficiency wedges. CKM also discuss several other studies which propose micro-level distortions that will show up as efficiency wedges. In short, CKM argue that the promising avenue for research is to write down economies in which primitive shocks in a detailed economy end up acting equivalently to productivity shocks in the prototype growth model.

Not Mismeasurement

Note that the CKM proposal is completely at odds with the view that measured productivity shocks are really just measurement error. We think there is a large amount of confusion on this point, and so we will elaborate.

Specifically, one well-known line of research argues that capital utilization varies significantly over the cycle and that if this variation is taken into account, then the size of the measured productivity shock is greatly diminished. This line concludes that therefore the size of the output fluctuations that can be accounted for by these shocks is diminished by a similar magnitude. A moment's reflection and some simple algebra show that while the premise may be true, the conclusion does not logically follow. (See CKM Proposition 3 for details.)

To make our point forcefully, we consider an extreme view about the variability in capital utilization. To accommodate this view, assume that the production function is now

(6) $$y = A(kh)^{\alpha}(nh)^{1-\alpha},$$

where n is the number of workers employed and h is the length (or *hours*) of the workweek. The labor input is, then, $l = nh$. Suppose that in the data, we measure only the labor input l and the capital stock k. We do not directly measure h or n. Our earlier specification can be interpreted as assuming that all of the observed variation in measured labor input l is in the number of workers and that the workweek h is constant. Under this interpretation, our benchmark specification, which we will refer to as the *fixed capital utilization* specification, correctly measures the efficiency wedge (up to the constant h).

Now consider the opposite extreme: assume that the number of workers n is constant and that all the variation in labor is from the workweek h. Under this *variable capital utilization* specification, the services of capital kh are proportional to the product of the stock k and the labor input l, so that variations in the labor input induce variations in the flow of capital services, and we can write the production function as $y = k^{\alpha}l$. We view our variable utilization specification as an extreme, because it implies that labor uses capital in fixed proportions so that as a given percentage of labor is laid off, the same percentage of capital is not utilized.[1] Presumably, a reasonable specification lies somewhere between these extremes.

In Figure 3, we plot the measured efficiency wedges for these two specifications of capital utilization during the Great Depression period (with it *fixed* in the benchmark economy and *variable* now). Clearly, when capital utilization is variable rather than fixed, the efficiency wedge falls less. Indeed, from 1929 to 1933, it falls only about half as much under the variable utilization specification as it does under the fixed utilization specification (about 9% rather than the previous 18%). But, as Figure 4 shows, the conclusion that the fraction of the output fall accounted for by the efficiency wedge also is cut in half is demonstrably false. Indeed, output falls by similar magnitudes from 1929 to 1933 in the two economies. What is going on is that while the change from the fixed to the variable capital utilization specification made the measured efficiency wedge fall, the change simultaneously made the equilibrium output response more sensitive to a given fall in that wedge, and these two effects effectively canceled each other out.

B. The Labor Wedge: Potential Detailed Models

We briefly discuss some detailed models that can give rise to the labor wedge in a prototype economy. This discussion is useful in two respects. It helps focus attention on particular promising models of the labor wedge. It also sets up our discussion of possible interpretations of the markup shock in Smets and Wouters' (2007) model.

One such model we find promising has government policies toward unions fluctuate. Another model, which we find more mechanical, has the consumer's value of leisure fluctuate. A third model, which we also find promising, is one with sticky wages. Here we show how the first two models can give rise to labor wedges; CKM does that for sticky wage models.

Fluctuating Government Policy Toward Unions

Consider, then, the following economy in which fluctuations in policies toward unions show up as fluctuations in labor market distortions in the prototype economy. (See Cole and

[1] While this may make sense for some forms of equipment, it clearly does not make sense for all of them: If 70% of the normal staff of a railroad show up to work, we don't need to leave 30% of the train behind. Or when farmers lay off 50% of their temporary farmhands at harvest, they aren't forced to drive their tractors only 50% of the time.

Ohanian (2004) for a discussion of such policies during the Great Depression.)

The technology for producing final goods from capital and a labor aggregate at history s^t is constant returns to scale and is given by

$$(7) \qquad y(s^t) = F\big(k(s^{t-1}), l(s^t)\big),$$

where $y(s^t)$ is output of the final good, $k(s^{t-1})$ is capital, and

$$(8) \qquad l(s^t) = \left[\int_0^1 l(i, s^t)^{\frac{1}{1+\lambda}} di\right]^{1+\lambda}$$

is an aggregate of the differentiated types of labor $l(i, s^t)$. Capital is accumulated according to (1). The problem faced by the final goods producer is to

$$(9) \qquad \max \sum_{t=0}^{\infty} \sum_{s^t} q(s^t) \left[y(s^t) - x(s^t) - w(s^t)l(s^t)\right],$$

where $q(s^t)$ is the price of a unit of consumption goods at s^t in an abstract unit of account and $w(s^t)$ is the aggregate real wage at s^t. The producer's problem can be stated in two parts. First, the producer chooses sequences for capital $k(s^{t-1})$, investment $x(s^t)$, and aggregate labor $l(s^t)$ subject to (7) and (1). Second, the demand for labor of type i by the final goods producer is

$$(10) \qquad l^d(i, s^t) = \left(\frac{w(s^t)}{w(i, s^t)}\right)^{\frac{1+\lambda}{\lambda}} l(s^t),$$

where $w(s^t) \equiv \left[\int w(i, s^t)^{-\frac{1}{\lambda}} di\right]^{-\lambda}$ is the aggregate wage.

There is a representative union that, when setting its wage, faces a downward-sloping demand for its type of labor, given by (10). The problem of the ith union is to maximize

$$(11) \qquad \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \pi(s^t)\, u\left(c(i, s^t), l(i, s^t)\right)$$

subject to the budget constraints

$$c(i, s^t) + \sum_{s_{t+1}} q(s^{t+1}|s^t) b(i, s^{t+1}) \leq w(s^t) l^d(i, s^t) + b(i, s^t) + d(s^t)$$

and the borrowing constraint $b(s^{t+1}) \geq -\bar{b}$, where $l^d(i, s^t)$ is given by (10).

Here $b(i, s^t, s_{t+1})$ denotes the consumers' holdings of one-period state-contingent bonds purchased in period t and state s^t, with payoffs contingent on some particular state s_{t+1} at

$t{+}1$, and $q(s^{t+1}|s^t)$ is the bonds' corresponding price. Clearly, $q(s^{t+1}|s^t) = q(s^{t+1})/q(s^t)$. Also, $d(s^t) = y(s^t) - x(s^t) - w(s^t)l(s^t)$ are the dividends paid by the firms. The initial conditions $b(i, s^0)$ are given and assumed to be the same for all i.

The only distorted first-order condition for this problem is that

$$(12) \quad w(i, s^t) = (1 + \lambda)\frac{u_l(i, s^t)}{u_c(i, s^t)}.$$

Notice that real wages are set as a markup over the marginal rate of substitution between labor and consumption. Clearly, given the symmetry among the consumers, we know that all of them choose the same consumption, labor, bond holdings, and wages, which we denote by $c(s^t)$, $l(s^t)$, $b(s^{t+1})$, and $w(s^t)$, and the resource constraint is as in (2).

We think of government pro-competitive policy as limiting the monopoly power of unions by pressuring them to limit their anti-competitive behavior. We model the government policy as enforcing provisions that make the unions price competitively if the markups exceed, say, $\bar{\lambda}(s^t)$, where $\bar{\lambda}(s^t) \leq \lambda$. Under such a policy, then, the markup charged by unions is $\bar{\lambda}(s^t)$, so that the key distorted first-order condition is that

$$(13) \quad w(s^t) = [1 + \bar{\lambda}(s^t)]\frac{u_l(s^t)}{u_c(s^t)}.$$

We now show that this detailed economy has aggregate allocations which coincide with those in a prototype economy. In that prototype economy, the firm maximizes the present discounted value of dividends

$$(14) \quad \max \sum_{t=0}^{\infty}\sum_{s^t} q(s^t)\left[F\big(k(s^{t-1}), l(s^t)\big) - x(s^t) - w(s^t)l(s^t)\right]$$

subject to $k(s^t) = (1 - \delta)k(s^{t-1}) + x(s^t)$. Consumers maximize

$$(15) \quad \sum_{t=0}^{\infty}\sum_{s^t} \beta^t \pi(s^t)\, u\left(c(s^t), l(s^t)\right)$$

subject to

$$(16) \quad c(s^t) + \sum_{s_{t+1}} q(s^{t+1}|s^t)b(s^{t+1}) \leq [1 - \tau(s^t)]w(s^t)l(s^t) + b(s^t) + d(s^t) + T(s^t),$$

where the dividends $d(s^t) = F\big(k(s^{t-1}), l(s^t)\big) - x(s^t) - w(s^t)l(s^t)$ and the lump-sum transfers $T(s^t) = \tau(s^t)w(s^t)l(s^t)$. The resource constraint is as in (2). The only distorted first-order condition is that

$$(17) \quad [1 - \tau(s^t)]w(s^t) = \frac{u_l(s^t)}{u_c(s^t)}.$$

Comparing (13) and (17), we see that the following proposition immediately follows:

Proposition 1. Consider the prototype economy just described with the following stochastic process for labor wedges:

$$(18) \quad 1 - \tau(s^t) = \frac{1}{1 + \bar{\lambda}(s^t)}.$$

The equilibrium allocations and prices of this prototype economy coincide with those of the unionized economy.

Fluctuating Utility of Leisure

In the detailed economy, let consumers' discounted utility be of the form (15), where the period utility function is separable and of the form

$$(19) \quad u\left(c(s^t), l(s^t)\right) = u\left(c(s^t)) + \psi(s^t)v(l(s^t)\right),$$

where $\psi(s^t)$ is an exogenous stochastic shock to the utility of leisure. The consumer maximizes utility (15) subject to the budget constraint

$$c(s^t) + \sum_{s_{t+1}} q(s^{t+1}|s^t)b(s^{t+1}) \leq w(s^t)l(s^t) + b(s^t).$$

The firm's problem here is identical to that in (14). The consumer's first-order condition for labor in this detailed economy is given by

$$(20) \quad \frac{v'(l(s^t))}{u'(c(s^t))} = \frac{w(s^t)}{\psi(s^t)}.$$

The associated prototype economy is nearly identical to the one described above. The consumer maximizes (15) subject to (16), where now the period utility function is of the form

$$(21) \quad u\left(c(s^t), l(s^t)\right) = u\left(c(s^t)) + v(l(s^t)\right),$$

which is the same separable form as in (19) except there is now no shock to the utility of leisure. The firm maximizes profits of the form (14). The consumer's first-order condition in this prototype economy is that

$$\frac{v'(l(s^t))}{u'(c(s^t))} = [1 - \tau(s^t)]w(s^t).$$

The following proposition is then immediate:

Proposition 2. Consider the prototype economy just described with the following stochastic process for labor wedges:

$$(22) \quad 1 - \tau(s^t) = \frac{1}{\psi(s^t)}.$$

The equilibrium allocations and prices of this prototype economy coincide with those of the detailed economy with a fluctuating value of leisure.

3. New Keynesian Models

The promise of New Keynesian models is twofold—first, that these models represent detailed economies that can generate the type of wedges we see in the data from interpretable primitive shocks; and second, that these models have enough microfoundations that both their shocks and parameters are *structural*, in that they can reasonably be argued to be invariant to monetary policy shocks. A model with both of these features would potentially be useful for monetary policy analysis. Unfortunately, the New Keynesian models are not. These models cannot generate the type of wedges we see in the data from interpretable primitive shocks. And it is doubtful that many of the features added on in the quantitative implementation of the models are structural. Hence, the models are not yet useful for policy analysis.

Here we focus on the Smets and Wouters (2007) model because it is widely considered the state-of-the-art New Keynesian model. Indeed, a version of it is now being used at the European Central Bank to help inform policymaking. The Smets-Wouters model has seven exogenous random variables. We divide these into two groups. The *potentially structural shocks* group includes total factor productivity, investment-specific technology, and monetary policy. The *dubiously structural shocks* group includes wage markups, price markups, exogenous spending, and risk premia.

A. The Wage-Markup Shock—A Fancy Name for a Labor Wedge?

In the Smets-Wouters model, one shock, the *wage-markup shock*, accounts for the bulk of fluctuations in aggregates. This shock appears as an additive shock in a linearized wage equation that relates current wages to past and expected future wages. We argue that this shock is a dubiously structural reduced-form shock that mechanically plays exactly the same role as our labor wedge. Because its dominant shock is a reduced-form shock, the Smets-

Wouters model, as it stands, cannot be used for policy analysis without taking a stand on what this shock represents. We offer two interpretations: it could stand in for either fluctuations in workers' bargaining power or shocks to leisure. These interpretations have radically different implications for policy. Obviously, then, until we have concrete microevidence in favor of at least one of these interpretations, the New Keynesian model should not be used for policy analysis.

The additive shock to the linearized wage equation in the Smets-Wouters model is motivated as coming from shocks to the labor aggregator. This labor aggregator relates aggregate labor l_t to a continuum of differentiated types of labor services $l_t(i)$ according to

$$(23) \quad 1 = \int_0^1 \left(G \frac{l_t(i)}{l_t}; \lambda_t \right) \, di,$$

where λ_t is referred to as the *wage-markup* shock. For intuition's sake, we find it useful to focus discussion on a special case of this aggregator, the *constant-elasticity of substitution* case explored by Smets and Wouters (2003), in which $G(l_t(i)/l_t; \lambda_t) = (l_t(i)/l_t)^{\frac{1}{1+\lambda_t}}$, so that

$$(24) \quad l_t = \left[\int_0^1 l_t(i)^{\frac{1}{1+\lambda_t}} di \right]^{1+\lambda_t}.$$

Clearly, making λ_t stochastic is just a simple way to make stochastic the elasticity of substitution between different types of labor in the labor aggregator (24), namely, $(1 + \lambda_t)/\lambda_t$.

We begin by showing that most of the movements in labor in the Smets-Wouters model are accounted for by this wage-markup shock. Using the estimated model, we can back out a time series for aggregate labor when the only stochastic shock is the wage-markup shock, denoted $l_t(\lambda)$. In Figure 5, we plot this series along with actual labor in the U.S. data.[2] Clearly, the bulk of the movements in labor are accounted for by this shock. For example, in a variance decomposition of labor, we find that two-thirds of labor's variance is accounted for by this shock. Not only does the wage-markup shock account for the bulk of the movements in labor, it accounts for about two-thirds of the variance of consumption and about half of the variance of output.

Given our business cycle accounting analysis, we are not surprised that this wage-markup shock plays a central role in generating fluctuations. We argue that this shock is

[2] Labor in the U.S. data is measured as total hours worked per person in the nonfarm business sector multiplied by the total number of civilians employed (workers aged age 16 years and older).

equivalent to a labor wedge. To see this equivalence, consider a stripped down flexible-wage version of the Smets-Wouters model with period utility function $u(c_t, l_t)$. Here, as in our 2002 interpretation, think of consumers as being organized into unions, so that the ith union consists of all consumers with labor of type i. The first-order condition for union i is to set the nominal wage for that type of labor $W_t(i)$ so that the corresponding real wage $w_t(i) = W_t(i)/P_t$ satisfies $w_t(i) = (1+\lambda_t)u_{lt}/u_{ct}$. Since all unions are symmetric, $w_t(i)$ equals the aggregate real wage w_t. This model therefore implies that

$$(25) \quad w_t = (1+\lambda_t)\frac{u_{lt}}{u_{ct}}.$$

(If we also abstract from sticky prices and monopoly power by firms, both of which play a quantitatively minor role in generating fluctuations in labor in the Smets-Wouters model, we have that the real wage equals the marginal product of labor.)

Now compare the wedge between the real wage and the marginal utility of leisure in (25) to the corresponding wedges in the two models described earlier and characterized by equations (18) and (22) of Propositions 1 and 2. Clearly, all the wage-markup shock λ_t does is generate a labor wedge in the model. In this sense, adding this shock is completely equivalent to mechanically sticking into the model an exogenous labor wedge, as we did in the prototype model.

We have already argued that the wedges identified in business cycle accounting cannot, by themselves, be used for policy analysis. Can the wage-markup shock? Consider a literal interpretation in which the wage-markup shock consists of fluctuations in the elasticity of substitution for different types of labor. To help with interpretation of units, we consider the constant-elasticity of substitution case with the labor aggregates given by (24). We re-estimated the Smets-Wouters model for this case after imposing, as Smets-Wouters did, that the mean markup was 50%. We found that the standard deviation of the markup was absurdly large, 2,587%. In the Smets-Wouters model, fluctuations in λ_t, taken literally, correspond to fluctuations in the elasticity of substitution $\big((1+\lambda_t)/\lambda_t\big)$ between carpenters, plumbers, neurosurgeons, and the like. We take it as a given that everyone, including Smets and Wouters, would regard these fluctuations as being several orders of magnitude outside of a reasonable range. Hence, a literal interpretation of the wage-markup shock is not palatable.

We view it instead as a reduced-form shock that stands in for some deeper shocks.

Since the wage-markup shock accounts for the bulk of fluctuations, the Smets-Wouters model cannot be used for policy analysis until we take a stand on the deeper shocks that it represents. Specifically, we need to argue that the shock is invariant to monetary policy. Furthermore, this shock must be interpretable enough so that we know whether it is a "bad shock," which policy should seek to offset, or a "good shock," which policy should seek to accommodate.

We turn now to two interpretations of the wage-markup shock.

Bargaining Power of Unions

One possible interpretation of the wage-markup shock is that it represents the bargaining power of unions, in particular, and labor, more generally. What gives rise to the shock's fluctuations and are these shocks invariant to monetary policy? Those questions, of course, are impossible to answer given how reduced-form the model. We tend to doubt, however, that they are invariant to policy. Presumably, though, advocates of this view see the bargaining power of unions relative to firms as related to the outside opportunities of the union members and firms. The whole point of a monetary policy intervention is to affect the real side of the economy and thus to change these opportunities. So this interpretation fails the policy-invariant requirement.

For argument's sake, however, suppose we do view these shocks as standing in for fluctuations in bargaining power and invariant to monetary policy interventions. The question then is, do we end up with a view of business cycles that most macroeconomists would find appealing? Under this interpretation, fluctuations in the bargaining power of workers lead them to become discontent at working at their current wages and to try to bid wages up. If workers are unsuccessful at bidding up their wages, they quit (so as to satisfy (25)), and if they are successful, the firm lays them off. Of course, if the model is to be consistent with the fact that wages are not countercyclical in the data, then what must be happening is that workers attempt to bid up their wages, fail to do so, become discontent, and quit. Hence, in equilibrium, the workers' greediness for higher wages simply leads to a fall in both their real income and their utility.

Under this interpretation, fluctuations in this shock are "bad," and the government should use all of its powers to offset their real effects on the economy. Indeed, the general principle here is that policy should be set so as to replicate the efficient equilibrium in which there is no monopoly power by workers and no sticky wages. In this efficient equilibrium, all variables, including labor, are at their efficient levels. Since most of the movements in labor are driven by this wage-markup shock, it will not be volatile. Monetary policy, which is a very poor tool for offsetting these shocks, should balance the benefits of keeping nominal wages constant against the other costs in the model of doing so.

Of course, if one actually believes that this type of shock drives the business cycle, then there is a much more powerful and effective policy to combat them: the government should crack down on unions very hard at the first hint of recession. Such a policy, which would be of the form that led to (13), would effectively eliminate business cycles in the U.S. economy.

Is this a palatable story of business cycles? We find it far-fetched to think that most New Keynesians would agree either that this is sensible policy or that it could eliminate most of the business cycle movement in labor. If, somehow, New Keynesians believe that worker greediness is responsible for recessions, then they should support this view with some detailed microeconomic evidence. For example, what fraction of the labor's fall in the recession can be accounted for by strikes?

The Value of Leisure

An alternative interpretation of the wage-markup shock is that it simply reflects consumers' utility of leisure along the lines discussed above. This interpretation of the shocks turns out to lead to an observationally equivalent economy in terms of aggregates to the one just discussed, but with vastly different policy implications. Thus, without more to go on than aggregate data, the policy implications of the New Keynesian model cannot even be pinned down. This finding is troubling to say the least.

To get some intuition for this obvservational equivalence result, consider an economy with a utility function of the form (19). Comparing (20) and (25), we see that in an economy in which the coefficient on leisure is given by

(26) $\psi(s^t) = 1 + \lambda(s^t),$

which has no distortions or monopoly power, the first-order condition for leisure will be equivalent to those in a stripped down flexible price version of the Smets-Wouters model with the fluctuations in monopoly power that gave rise to (25).

The Smets-Wouters model is actually more complicated than the stripped-down version because with the Calvo-type way of making wages sticky, wages are set as a markup over a present value of the marginal utility of leisure. But the equivalence between fluctuations in the value of leisure and fluctuations in monopoly power holds even in this setting. Indeed, as Smets and Wouters (2003, 2007) argue, in the log-linearized model they use in estimation, it is impossible to identify whether their wage-markup shocks are really shocks to the elasticity of substitution in the labor aggregator, as in (24), or shocks to leisure, as in (19).

Note that the policy implications of interpreting the wage-markup shock as fluctuations in leisure are radically different than those of the bargaining power interpretation. Under the leisure interpretation, fluctuations in the shock are "good," and the Fed should accommodate them. But this interpretation of the shock in the New Keynesian model has serious issues. To get a feel for these issues quantitatively, we followed Smets and Wouters (2003) and allowed for an AR(1) taste shock and an i.i.d markup shock (as did Levin et al. (2006)). We refer to this model as the *taste-shock* version of the Smets and Wouters model. In Figure 6, we plot the potential and actual output from 1965 to 2005 from the taste-shock version of the Smets-Wouters model estimated for the United States.

We see that in the period from 1979 to 1984, the United States went through two recessions that many economists attributed in good part to the Fed's actions aimed at reducing inflation. The figure shows that as output fell, so did output in the efficient equilibrium. Indeed, in much of the early 1980s, the efficient output level was lower than the observed output level.

In short, are the New Keynesians willing to accept their model's implication that the driving force behind the postwar recessions is that, in Modigliani's (1977) terminology, workers suffered contagious attacks of laziness? Are they willing to accept their model's implication that the recessions between 1979 and 1984 had almost nothing to do with monetary policy? Do they accept their model's implication that the Fed should have tightened even more during recessions because its actual monetary policy discouraged workers from taking the even longer vacations from working that they desired?[3]

In sum, we have difficulties with both interpretations of the key shock in the New Keynesian model and the associated policy recommendations. Presumably, the New Keynesians do as well.

B. A Dubious Mechanism for Generating Persistent Inflation

Consider next another feature of the New Keynesian model that has important implications for policy but has only a dubious structural interpretation.

Several researchers, including Fuhrer (1996) and Mankiw (2001), have pointed out that the simple New Keynesian models, even with Calvo price- and wage-setting, cannot generate persistent inflation. Motivated by some VAR evidence showing that inflation is persistent, Christiano, Eichenbaum, and Vigfusson (2004) have shown that by adding backward indexation of prices, the New Keynesian model can generate persistence in inflation. This, however, is a costly way to get this result.

Smets and Wouters (2003, 2007), building on the work of Christiano, Eichenbaum, and Vigfusson, incorporate this feature into their models. Specifically, Christiano, Eichenbaum, and Vigfusson assume that even those firms that are not allowed to freely adjust their prices at t, mechanically adjust them to lagged inflation, so that the price p_{jt} charged by a non-adjusting firm j in time period t equals

$$(27) \quad p_{jt} = \pi_{t-1} p_{jt-1},$$

where p_{jt-1} is this firm's price in $t-1$ and π_{t-1} is the rate of gross inflation of the aggregate price level between periods $t-1$ and t. Smets and Wouters (2003, 2007) assume something similar, except they allow for only partial indexation.

The problem with this assumption is that it is counterfactual. We know this thanks to the work of Bils and Klenow (2004), Midrigan (2006), Golosov and Lucas (2007), Nakamura

[3] Walsh (2006) expresses similar skepticism about this version of the New Keynesian model.

and Steinsson (2007), and others. Their evidence on price behavior at the micro level strongly suggests that the backward price indexing assumption is wrong.

To see this, consider the actual prices charged for a particular product in scanner data from a grocery store. In Figure 7, we plot the price charged for a package of Angel Soft Bathroom tissue at Dominick's Finer Food retail store in Chicago along with what the price would look like if it were backward-indexed along the lines of (27) as is assumed by Christiano, Eichenbaum, and Evans (2005). Clearly, the path of the actual price does not look like that assumed. We have picked on particular series to illustrate our point but we could have shown literally thousands more that look similar.

More generally, the key statistics reported in the budding literature on the properties of individual prices are the average number of months before the price is changed. Bils and Klenow (2004) report that number to be on the order of four months, while Nakamura and Steinsson (2007) use a different procedure and report a number on the order of eleven months. Note that the New Keynesian model's predictions are simply flatly inconsistent with the micro data. If we used either Bils and Klenow's algorithm or Nakamura and Steinsson's algorithm on prices generated from the New Keynesian models, we would find that prices changed every single period.

There seems to be some confusion on this point in the literature that uses the backward indexation assumption. When, for example, Bils and Klenow report that the average time between price changes is four months they are not providing an estimate of the Calvo probability of changing a price in a economy which, because of backward indexation, all prices change in every period. Rather Bils and Klenow's numbers imply that to be consistent with the micro data the model has to have the prices be completely and utterly fixed between price changes and then on average that price changes every four months.

Hence, while sticking an ad hoc backward price indexation equation of the form of (27) into a model can make the model mechanically generate persistence in inflation, the mechanism by which it does so is flatly inconsistent with the micro data.

Aside from that inconsistency, the problem with proceeding in this mechanical fashion is that the backward indexation feature shapes the policy advice from the model. In particular, as the literature has shown, the costs of disinflation in an economy with backward indexation are quite high. If the persistence of inflation is coming from another mechanism, then there may not be such high costs.

C. The Dubious Model of the Fed's Policy Function

The question naturally rises, is there a plausible mechanism that can generate the persistence in inflation that we see in the data in a way that is not inconsistent with the micro evidence? Yes. We argue that the persistence of inflation naturally arises from a random walk–like feature of interest rate policy that is being missed in the current model.

New Keynesian models assume that short-term nominal rates are stationary and ergodic; hence, the long-term nominal rates implied by that rule are much too smooth relative to the observed long-term nominal rates in the data. We argue that this discrepancy leads the New Keynesian models to misidentify the source of persistence in inflation, and hence, leads these models to give erroneous policy advice about the costs of disinflation.

The gist of our argument follows from two features of the data. First, as is well-known, during the postwar period, short rates and long rates have a very similar secular pattern. (For some recent work documenting this feature, see the 2008 work of Atkeson and Kehoe.) Second, a large body of work in finance has shown that the level of the long rate is well-accounted for by the expectations hypothesis. (See, for example, the 2008 work of Cochrane and Piazzesi.) Combining these two features of the data implies that when the Fed alters the current short rate, private agents significantly adjust their long-run expectations of the future short rate, say, 30 years into the future. At an intuitive level, then, we see that Fed policy has a large random walk component to it.

When we incorporate this persistent feature of policy into a model, the model naturally delivers persistence in inflation. Indeed, as Ireland (2007) shows, once we allow the Fed policy function to have a random walk component, the model needs no backward indexation of prices in order to fit the data. Indeed, if we run a horserace between two models—one with a standard Taylor rule and backward indexation and one with a random walk component to interest rate policy and no backward indexation—the second model fits the data better.

Under this view of research, what happened is the following. Because the standard New Keynesian model does not adequately incorporate the random walk component of policy that the data on long rates call out for, a simple version of the model without backward indexation does not generate enough persistence in inflation. (See Collard and Dellas (2005) for a demonstration.) To get the model to generate persistence, researchers have mechanically added backward indexation of prices (and wages). The model so constructed implies that disinflation is very costly. However, if we recognize that the persistence in inflation is coming from persistence in policy, then no backward indexation is needed, and this version of the model implies rather small costs from disinflation. In this sense, trying to fix an empirical problem by adding mechanical features makes the model give the wrong answer to a basic policy question.

D. Dubious Other Shocks

So far we have argued that wage markups are dubiously structural. Similar concerns apply to price markups. We now argue that also dubiously structural are the risk premium shocks and exogenous spending shocks also added to help New Keynesian models fit the data.

Consider the risk premium shocks. (By the way, we find the term *risk premium shocks* exceptionally confusing because the Smets-Wouters model has no risk premium.) These shocks enter the consumer's first-order condition for government debt, but not the first-order condition for accumulating capital. In this sense, these shocks resemble (unobserved) time-varying taxes on short-term nominal government debt (relative to taxes on capital income). In the Smets-Wouters model, these shocks are enormous. For example, they are 6.5 times as variable as short-term nominal interest rates.

The only sensible economic interpretation that we can give to these shocks is that they are meant to capture financial market episodes when there is a "flight to quality" in the sense that consumers' preference for holding government debt increases abruptly. Unfortunately for the Smets-Wouters model, under this interpretation, these shocks are hardly likely to be structural with respect to monetary policy.

Consider next the exogenous spending shocks. These shocks are referred to by Smets and Wouters as "government spending shocks." Unfortunately, the resulting shocks have little to do with measured government spending. For example, the correlation between government spending in the data (defined as the sum of government consumption and investment) and the

Smets-Wouters measure of government spending is about .39, and the variance of government spending in the data is more than 6 times the variance of the Smets and Wouters measure of government spending. The reason is that in the Smets-Wouters empirical implementation, these shocks are residually defined from the national income identity and include, among other variables, net exports. Variables like net exports are also not likely to be structural with respect to monetary policy.

4. Conclusion

New Keynesian models are not yet useful for policy analysis. The main reason is that model builders in this tradition have added so many free parameters that the features and shocks in their models are only dubiously structural.

Changes in method can make these models potentially useful for policy analysis. The most important change in method needed is to resist the urge to add undisciplined free parameters in order to fit the same old aggregate time series. A far preferable procedure is to start with a small model, add features and shocks, one at a time, carefully disciplined by appropriate microeconomic evidence.

One example, specifically set in the context of the Smets-Wouters model, is to begin by noting that this model has large fluctuations in the cross-sectional distribution of employment, fluctuations that are inefficient. The primary job of optimal monetary policy is to reduce fluctuations in the cross-sectional distribution of employment by reducing the cross-sectional distribution of wages over the business cycle. (See Levin et al. (2006).) Given the importance of these cross-sectional distributions for shaping monetary policy, at the very minimum, researchers pursuing variants of the Smets-Wouters model should ask whether the data show significant fluctuations in these distributions as well as the links between the cross-sectional distributions of wages and employment. If the data appear promising in this regard, then these data should be used to discipline the estimation. If the data are not promising, then it is best to look elsewhere for a model.

Processes of this kind will be slow and painful, but will avoid the false promise of the Old Keynesian revolution that the profession had trustworthy tools for designing and implementing good policy.

References

Atkeson, Andrew, and Patrick J. Kehoe. 2008. "On the Need for a New Approach to Analyzing Monetary Policy." Manuscript. Federal Reserve Bank of Minneapolis.

Bils, Mark, and Peter J. Klenow. 2004. "Some Evidence on the Importance of Sticky Prices." *Journal of Political Economy,* 112 (October): 947–85.

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan. 2002. "Can Sticky Prices Generate Volatile and Persistent Real Exchange Rates?" *Review of Economic Studies*, 69 (3, July): 533–63.

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan. 2007. "Business Cycle Accounting." *Econometrica,* 75 (3, May): 781–836.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy*, 113 (1, February): 1–45.

Christiano, Lawrence J., Martin Eichenbaum, and Robert Vigfusson. 2004. "What Happens After a Technology Shock?" Manuscript. Northwestern University.

Cochrane, John H., and Monika Piazzesi. 2008. "Decomposing the Yield Curve." Manuscript. University of Chicago.

Cole, Harold L., and Lee E. Ohanian. 2004. "New Deal Policies and the Persistence of the Great Depression: A General Equilibrium Analysis." *Journal of Political Economy*, 112 (4, August): 779–816.

Collard, Fabrice, and Harris Dellas. 2006. "Dissecting the New Keynesian Model." Manuscript. Toulouse School of Economics.

Fuhrer, Jeffrey C. 1996. "Monetary Policy Shifts and Long-Term Interest Rates." *Quarterly Journal of Economics*, 111 (4, November): 1183–1209.

Golosov, Mikhail, and Robert E. Lucas, Jr. 2007. "Menu Costs and Phillips Curves." *Journal of Political Economy*, 115 (April): 171–99.

Ireland, Peter N. 2007. "Changes in the Federal Reserve's Inflation Target: Causes and Consequences." *Journal of Money, Credit, and Banking*, 39 (8, December): 1851–82.

Kimball, Miles S. 1995. "The Quantitative Analytics of the Basic Neomonetarist Model." *Journal of Money, Credit, and Banking*, 27 (4): 1241–77.

Levin, Andrew T., Alexei Onatski, John C. Williams, and Noah M. Williams. 2006. "Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models." In *NBER Macroeconomics Annual 2005*, ed. Mark Gertler and Kenneth Rogoff, 229–87. Cambridge, MA: MIT Press.

Lucas, Robert E., Jr., and Nancy L. Stokey. 1983. "Optimal Fiscal and Monetary Policy in an Economy without Capital." *Journal of Monetary Economics*, 12 (July): 55–93.

Mankiw, N. Gregory. 2001. "The Inexorable and Mysterious Tradeoff between Inflation and Unemployment." *Economic Journal*, 111 (471, May): C45–61.

Midrigan, Virgiliu. 2006. "Menu Costs, Multi-Product Firms, and Aggregate Fluctuations." Manuscript. Federal Reserve Bank of Minneapolis.

Modigliani, Franco. 1977. "The Monetarist Controversy or Should We Foresake Stabilization Policies?" *American Economic Review*, 67 (2, March): 1–19.

Nakamura, Emi, and Jón Steinsson. 2007. "Five Facts About Prices: A Reevaluation of Menu Cost Models." Manuscript. Harvard University.

Smets, Frank, and Raf Wouters. 2003. "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association*, 1 (5, September): 1123–75.

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

Walsh, Carl E. 2006. "Comment on 'Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models.'" In *NBER Macroeconomics Annual 2005*, ed. Mark Gertler and Kenneth Rogoff, 297–308. Cambridge, MA: MIT Press.

Yun, Tack. 1996. "Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles." *Journal of Monetary Economics*, 37 (2, April): 345–70.

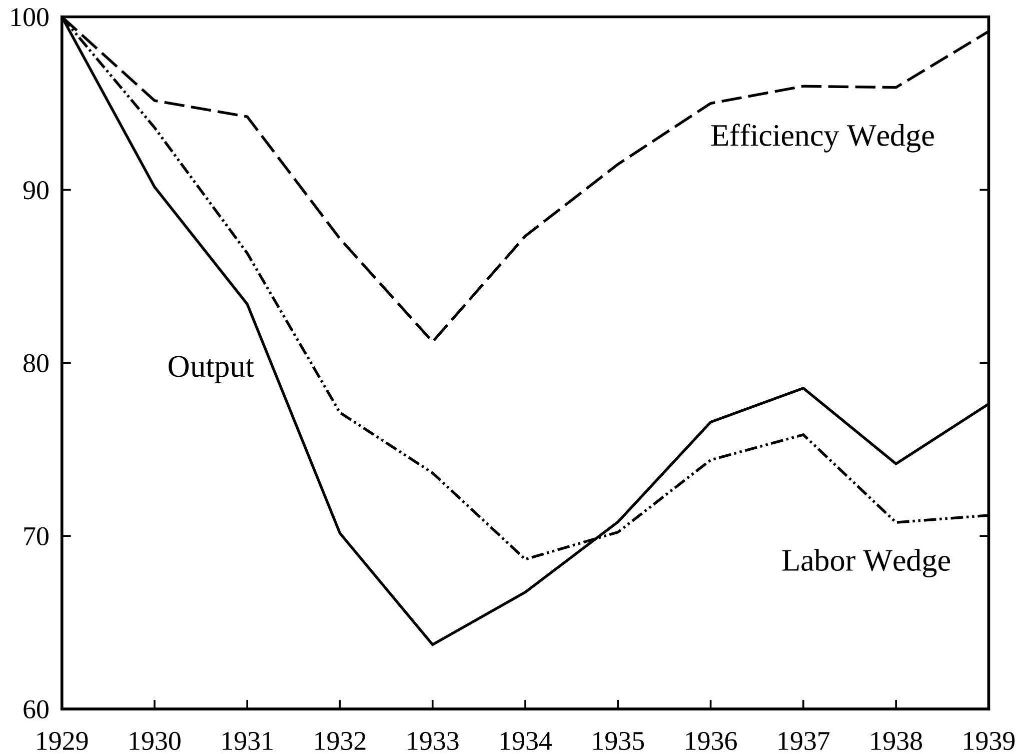

Figure 1. U.S. Output and Two Measured Wedges

Annual, Normalized to equal 100 in 1929

Source: Chari, Kehoe, and McGrattan (2007)

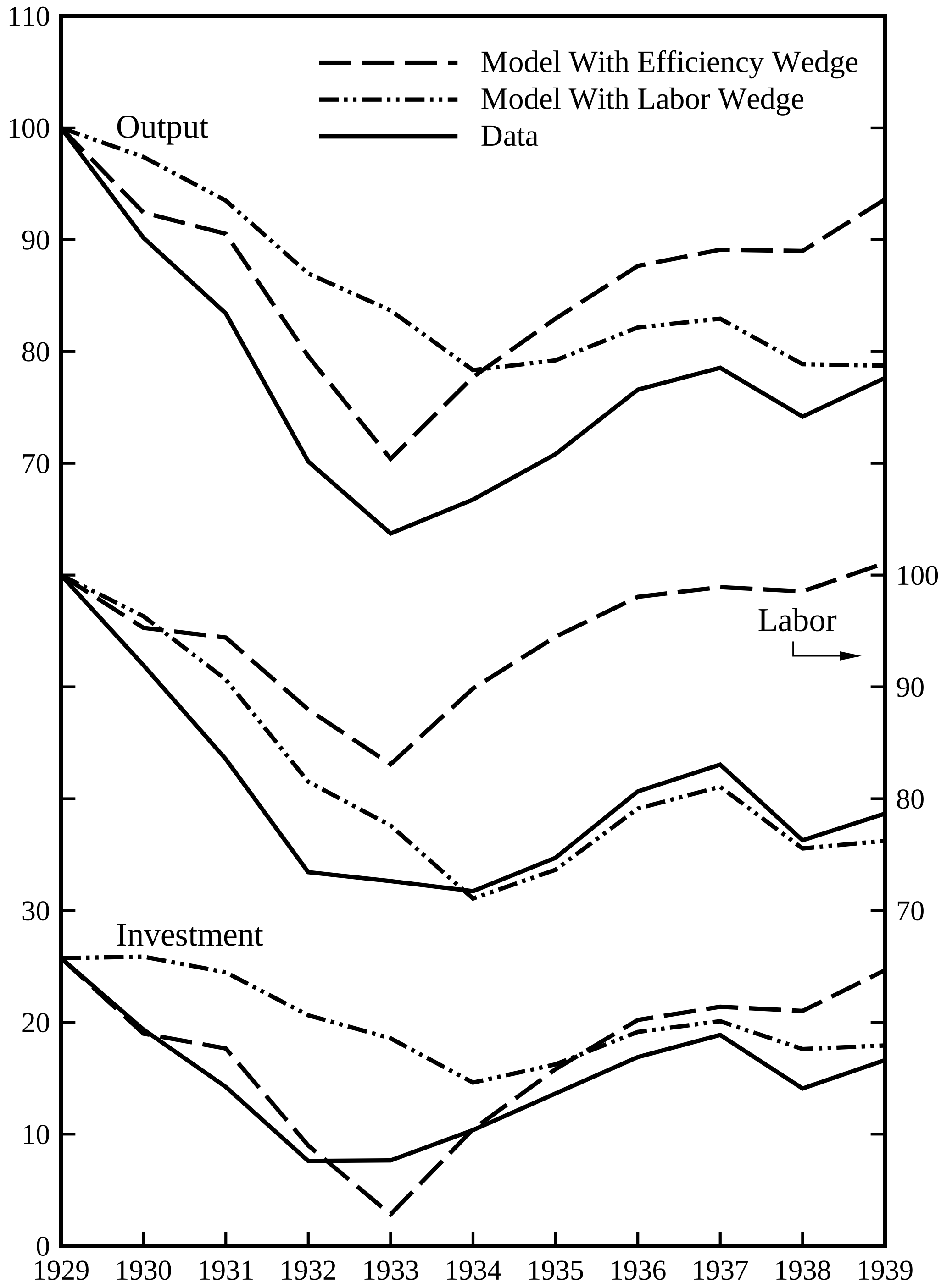

Figure 2. Data and Predictions of Models with Just One Wedge

Source: Chari, Kehoe, and McGrattan (2007)

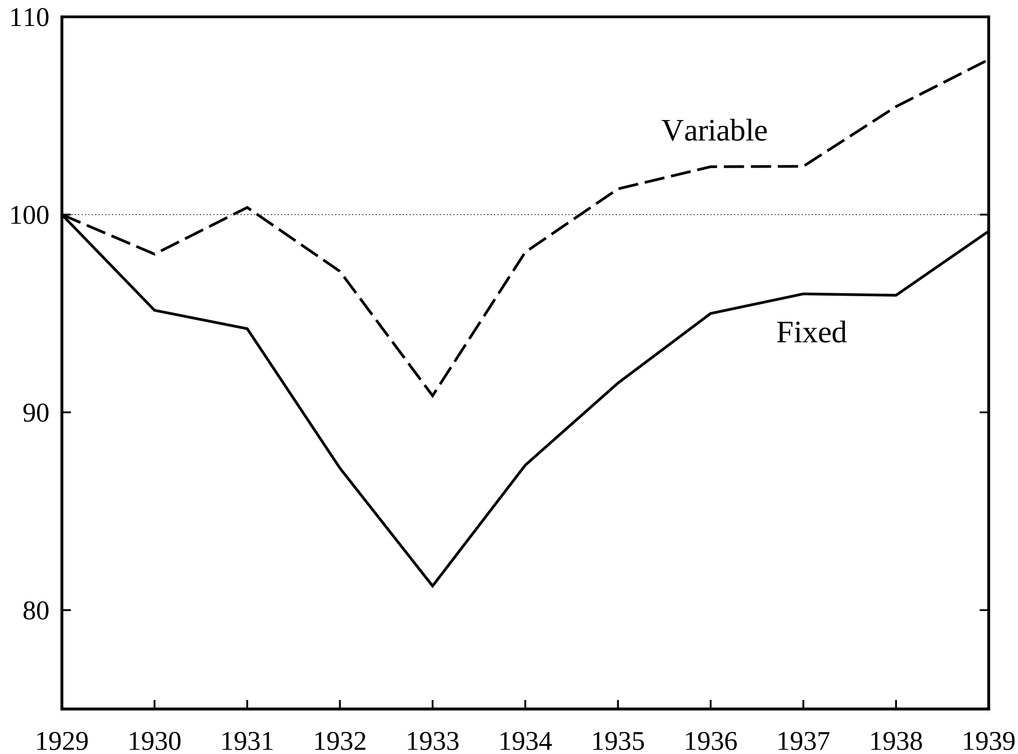

Figure 3. Measured Efficiency Wedges for Two Capital Utilization Specifications

Annual, Normalized to equal 100 in 1929

Source: Chari, Kehoe, and McGrattan (2007)

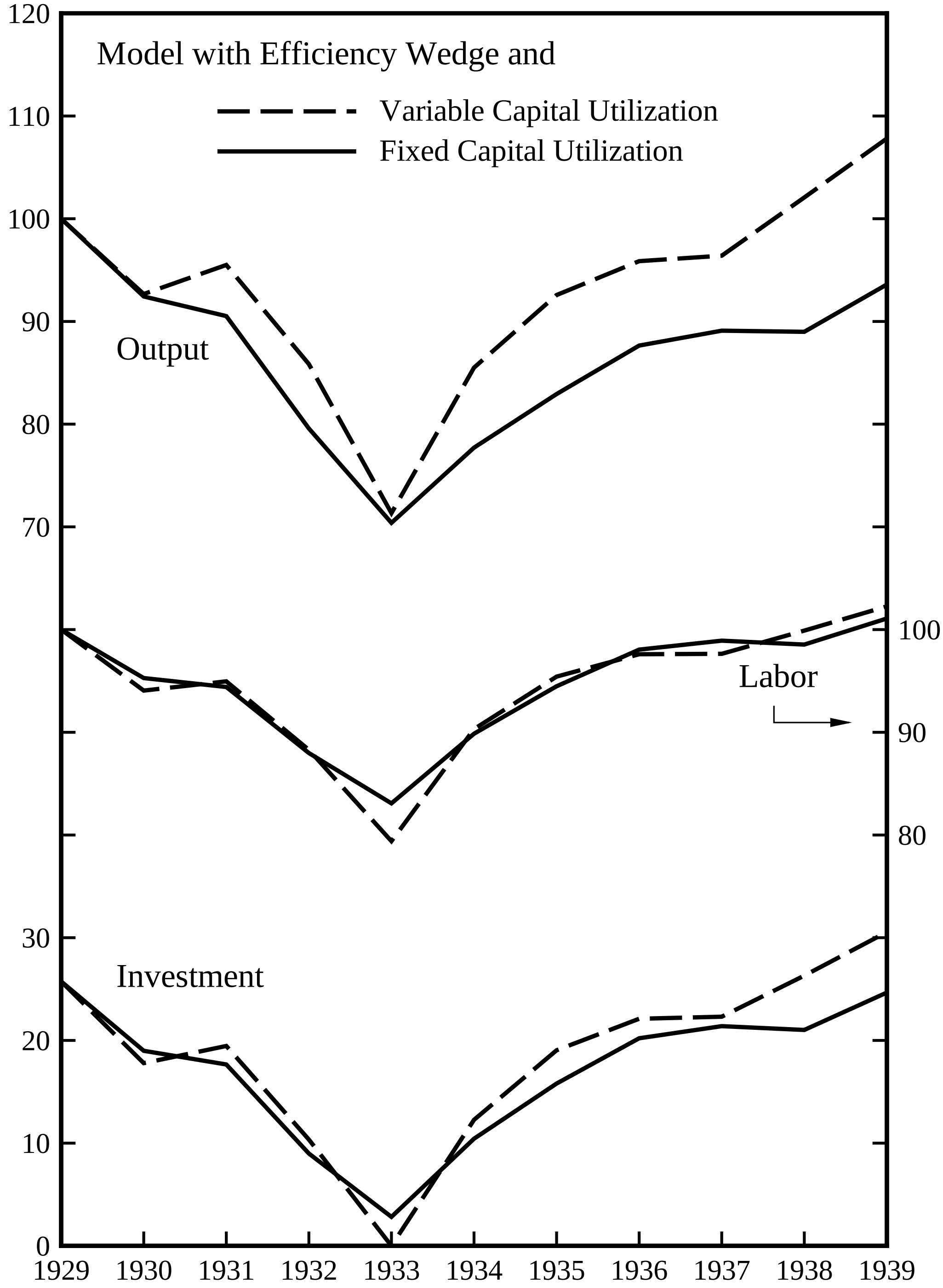

Figure 4. Data and Predictions of Models with Variable Capital Utilization and Only the Efficiency Wedge

Source: Chari, Kehoe, and McGrattan (2007)

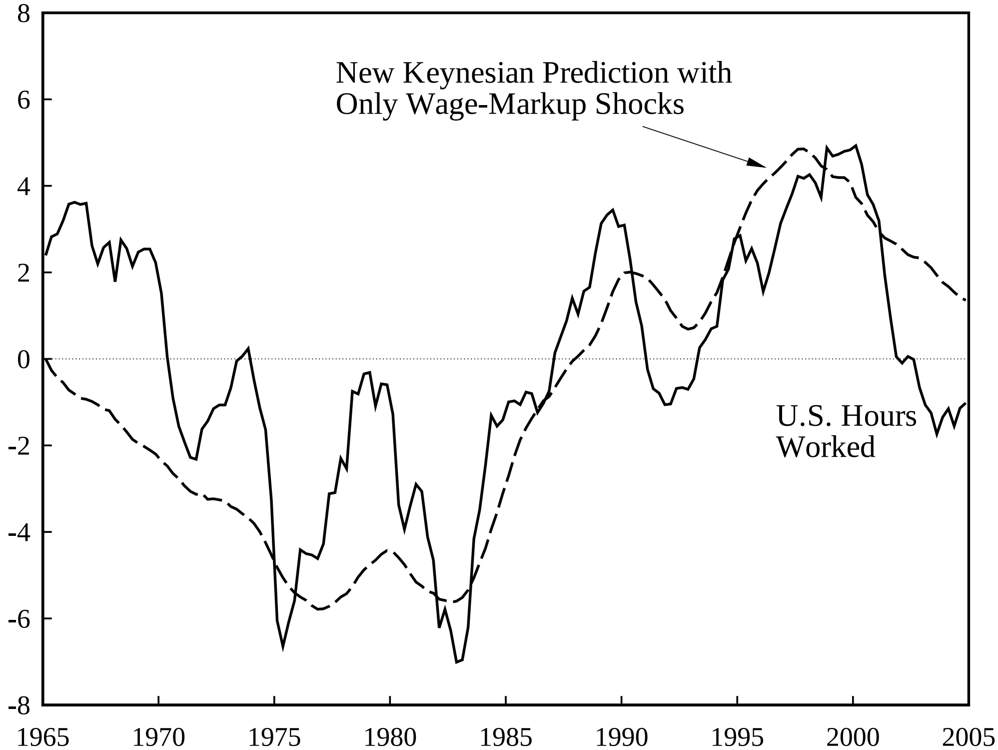

Figure 5. Data and Prediction of Smets and Wouters (2007) Model with Only Wage-markup Shocks

Source of U.S. Data: Bureau of Labor Statistics

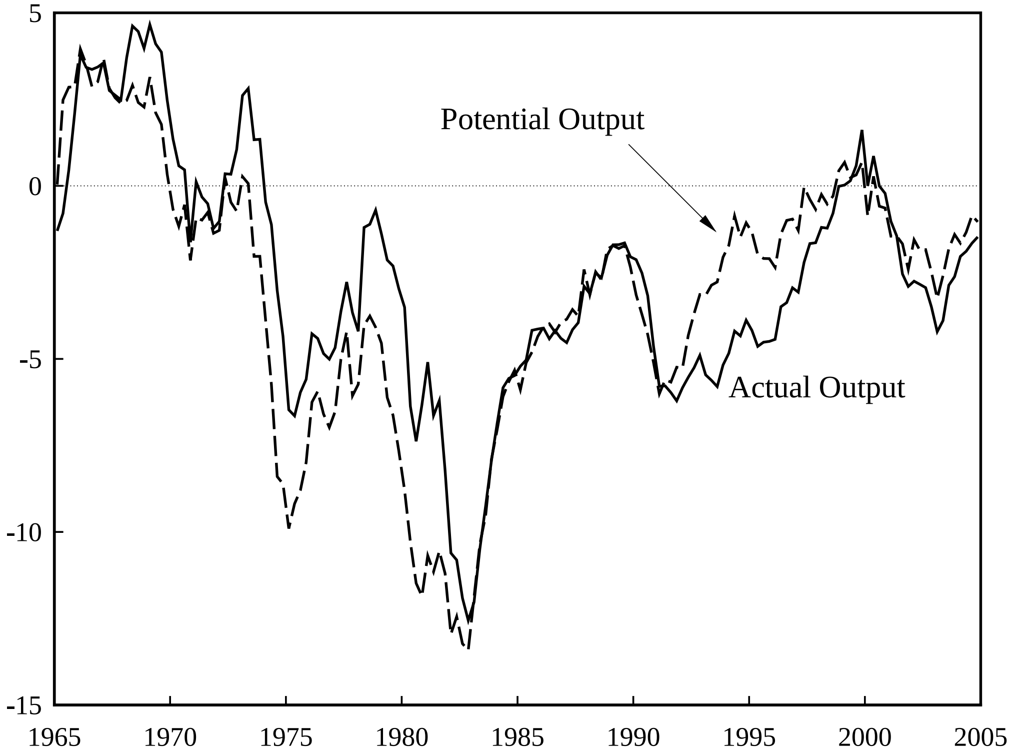

FIGURE 6. ACTUAL AND POTENTIAL OUTPUT IN VERSION OF SMETS-WOUTERS (2007) MODEL WITH AR(1) TASTE SHOCKS AND I.I.D. WAGE-MARKUP SHOCKS

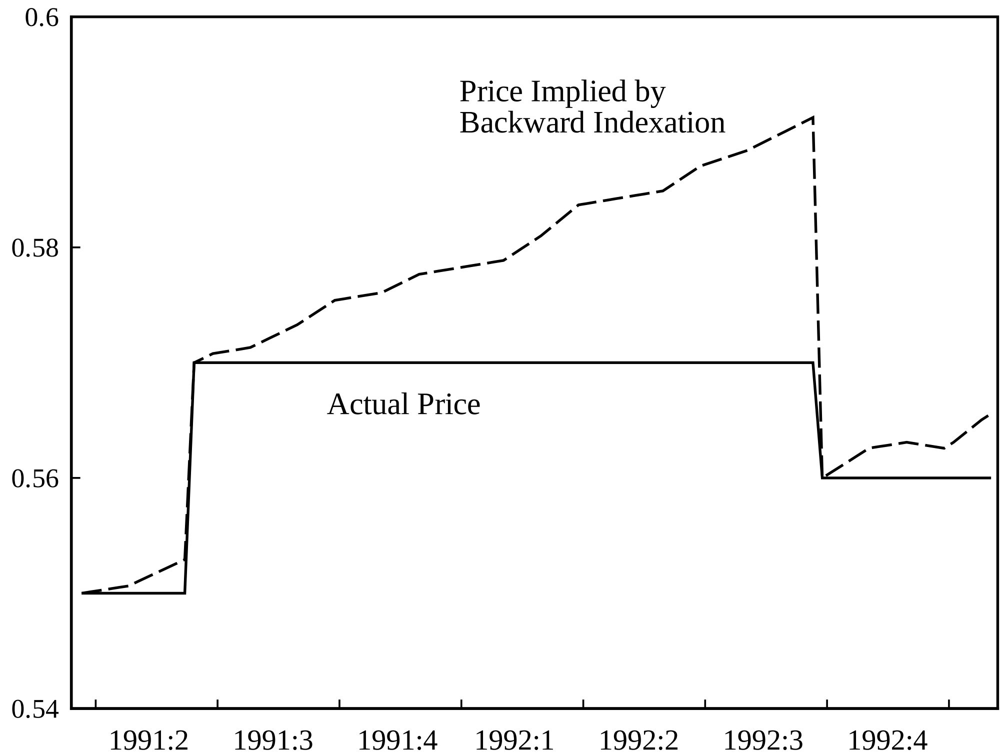

Figure 7. Weekly Price of Angel Soft Bathroom Tissue and Price Implied by Backward Indexation