Convergence in Macroeconomics: The Labor Wedge[†]

By Robert Shimer*

I review research on the behavior of the labor wedge, the ratio between the marginal rate of substitution of consumption for leisure and the marginal product of labor. According to competitive, market-clearing macroeconomic models, the ratio is easy to measure and should be equal to the sum of consumption and labor taxes. The observation that the wedge is higher in continental Europe than in the United States has proved useful for understanding the extent to which taxes can explain differences in labor market outcomes. The observation that the ratio rises during recessions suggests some failure of competitive, market-clearing macroeconomic models at business cycle frequencies. The latter observation has guided recent research, including work on sticky wage models and job search models. (JEL E24, E32, J64)

E conomics studies the interaction of optimizing households and firms in a marketplace. Modern macroeconomics is different from modern microeconomics only in its focus on aggregation and on general equilibrium. In particular, virtually all modern macroeconomic models build upon two foundations, regardless of whether they analyze business cycle fluctuations, study long-run growth, or make cross-country comparisons. First, households maximize expected utility subject to a budget constraint. Second, firms maximize expected profits. Of course, agreement that these two pieces should be elements of a macroeconomic model does not imply agreement on how the economy functions, on how it reacts to monetary or fiscal policy interventions, or on what optimal monetary and fiscal policy should look like. The answers to those questions, perhaps inevitably, depends on more controversial issues, including the nature of shocks hitting the economy and the formation and evolution of beliefs. Still, this essay argues that utility maximization and profit maximization have strong and testable implications for the patterns one should expect to see in the data. I will argue that these patterns have guided significant aspects of macroeconomics research in recent years and are likely to continue to do so in the future.

My focus is on the ratio of the marginal rate of substitution of consumption for leisure (MRS) and the marginal product of labor (MPL), the "labor wedge." Suppose

that each household believes that its choice of how much to consume and how much to work has no effect on its wage, each firm believes that its choice of how many workers to hire has no effect on the wage, the labor market clears, and there are no consumption or labor taxes. Then the MRS and MPL should be equal. In reality, these conditions are violated. For example, in any modern economy, taxes drive a wedge between the MRS and MPL. The first section of this essay reviews work by Edward C. Prescott (2004) (and others) that argues if labor supply is sufficiently elastic, the growing gap in labor and consumption taxes between the United States and many continental European countries can explain the growing gap in employment and hours worked. That is, changes in taxes drive the observed change in the wedge between the MRS and the MPL.

I then turn to business cycle frequency fluctuations. Viewed through the same lens, US data on consumption and labor supply indicate a counter-cyclical wedge between the MRS and the MPL. During recessions, workers and firms behave as if they face an increase in taxes on labor supply or consumption. In the absence of evidence of such an increase, I consider modifications to the basic model consistent with this empirical pattern. In recent years, the most common means of accounting for the labor wedge is through an assumption that the representative household's disutility of labor fluctuates at business cycle frequencies. Recessions are periods when households dislike working. An observationally equivalent hypothesis is that workers' wage setting power fluctuates at business cycle frequencies. Recessions are periods when households reduce their labor supply to drive up wages. I share the view of many economists that shocks to the disutility of labor for the representative household or to workers' wage-setting power are an unsatisfactory and probably misleading explanation of movements in the labor wedge. I finish the essay by arguing for a more promising, if still preliminary, explanation—search frictions, combined with real wage rigidities, create an endogenous cyclical wedge between the MRS and MPL.

This essay was prepared for a session on "Convergence in Macroeconomics," but I have decided to focus on one topic, the labor wedge. In part, this is a question of expertise. Macroeconomics is a broad field and I may not be qualified to comment on all aspects of it. But the labor wedge is also a natural focus for a session on convergence. As I highlight in this essay, it is central in the long run, where it determines the evolution of employment and hours worked, and in the short run, where understanding its behavior may be critical for explaining how shocks affect labor market outcomes. Thus, an analysis of the labor wedge highlights the fact that macroeconomists agree on the importance of using the same model to analyze the short run and the long run. Finally, most macroeconomists have converged on a model of the determination of the MRS and the MPL—household and firm optimization—and agree that there is significant variation in the measured labor wedge across countries and in the business cycle. The only controversial question is why it varies.

I. Cross-Country Analysis

My starting point is a simple model of the interaction between a representative household and a representative firm. I denote time by $t = 0, 1, 2, \ldots$ and the state

of the economy at time t by s_t. The current state includes variables such as aggregate productivity, government spending, and distortionary tax rates, determined outside the model. Let $s^t = \{s_0, s_1, \ldots, s_t\}$ denote the history of the economy and $\pi(s^t)$ denote the time 0 belief of the probability of observing an arbitrary history s^t through time t. Note that I do not take a stand on whether expectations are rational. For example, agents may believe at time 0 that the government debt will explode, even if this is inconsistent with the underlying model.

The household is infinitely lived and has preferences over history-s^t consumption $c(s^t)$ and history-s^t labor supply $h(s^t)$ ordered by

(1)
$$U(\{c, h\}) = \sum_{t=0}^{\infty} \beta^t \left(\sum_{s^t} \pi(s^t) \left(\log c(s^t) - \frac{\gamma \varepsilon}{1 + \varepsilon} h(s^t)^{\frac{1+\varepsilon}{\varepsilon}} \right) \right),$$

where $\beta \in (0, 1)$ is the discount factor, $\gamma > 0$ measures the disutility of working, and, as I show below, $\varepsilon > 0$ is the Frisch (constant marginal utility of wealth) elasticity of labor supply.

The household faces a sequence of intertemporal budget constraints:

(2)
$$a(s^t) + (1 - \tau_h(s^t)) w(s^t) h(s^t) + T(s^t) = (1 + \tau_c(s^t)) c(s^t)$$
$$+ \sum_{s_{t+1}} (1 + \tau_k(s^{t+1})) q(s^{t+1}) a(s^{t+1}),$$

where $s^{t+1} \equiv \{s^t, s_{t+1}\}$. The household starts a typical history s^t holding $a(s^t)$ units of a real bond. It then earns an after-tax wage $(1 - \tau_h(s^t)) w(s^t)$, where $\tau_h(s^t)$ is the labor income tax, and it receives a history-contingent, lump-sum transfer $T(s^t)$. It pays $1 + \tau_c(s^t)$ for each unit of consumption and pays $(1 + \tau_k(s^{t+1})) q(s^{t+1})$ to purchase a unit of assets in the continuation history s^{t+1}, where $\tau_k(s^{t+1})$ is the capital tax rate. In addition, the household must be able to pay off its debt following any history through an appropriate choice of consumption and labor supply, ruling out Ponzi games. This formulation allows for incomplete markets through constraints on assets. We can express market incompleteness through a vector of constraints taking the form $G(\{a\}) \geq 0$.[1]

The household chooses history-contingent consumption and hours subject to a sequence of budget constraints, the no-Ponzi game condition, and any additional restrictions on asset holdings. I focus on the first-order conditions with respect to history-s^t consumption and labor supply:

(3)
$$\frac{1}{c(s^t)} = \lambda(s^t)(1 + \tau_c(s^t)) \text{ and } \gamma h(s^t)^{1/\varepsilon} = \lambda(s^t)(1 - \tau_h(s^t)) w(s^t),$$

---

[1] For example, one can impose nonnegativity constraints, $a(s^t) \geq 0$, to prohibit borrowing. It is also possible to introduce constraints on the measurability of assets, e.g. $a(\{s^t, s_{t+1}\})$ is independent of s_{t+1}, which implies that only risk-free bonds can be traded.

where $\lambda(s^t)$ is the Lagrange multiplier on the history-s^t budget constraint, equation (2). Note, from the second equation, that a one percent increase in the after-tax wage $(1 - \tau_h)w$ raises labor supply h by ε percent, holding fixed the Lagrange multiplier λ. Thus, ε is the Frisch elasticity of labor supply, the key parameter in what follows. Eliminate $\lambda(s^t)$ between these equations and solve for the wage,

$$
(4) \qquad\qquad w(s^t) = \frac{\gamma c(s^t) h(s^t)^{1/\varepsilon}}{1 - \tau(s^t)},
$$

where $\tau(s^t) \equiv (\tau_c(s^t) + \tau_h(s^t))/(1 + \tau_c(s^t))$ is the relevant tax rate, a combination of the consumption and labor tax. An additional unit of pre-tax labor income in history-s^t permits $1 - \tau(s^t)$ additional units of consumption in that history, after paying the labor and consumption taxes. Equation (4) states that the wage is equal to the tax-adjusted MRS.

The representative firm has access to a Cobb-Douglas production function, producing net output $A(s^t)k^\alpha n^{1-\alpha} - \delta k$, where k and n are the capital and labor it employs, $A(s^t)$ is history-contingent total factor productivity, and δ is the depreciation rate. The firm rents capital at the interest rate $r(s^t)$ and pays workers the wage $w(s^t)$. For notational simplicity alone, I assume that households bear all the incidence of taxes. Then the firm solves

$$
(5) \qquad\qquad \max_{\{k,\,n\}} \; (A(s^t)k^\alpha n^{1-\alpha} - (\delta + r(s^t))\, k - w(s^t)n).
$$

The first-order condition for labor demand yields

$$
(6) \qquad\qquad w(s^t) = (1 - \alpha)y(s^t)/n(s^t),
$$

where $y(s^t) = A(s^t)k(s^t)^\alpha n(s^t)^{1-\alpha}$ is the firm's gross output. Equation (6) states that the wage is equal to the MPL.

Now eliminate the state-contingent wage between equations (4) and (6) and impose labor market clearing, $n(s^t) = h(s^t)$. Solving for $h(s^t)$ gives

$$
(7) \qquad\qquad h(s^t) = \left(\frac{(1 - \alpha)(1 - \tau(s^t))}{\gamma(c(s^t)/y(s^t))} \right)^{\frac{\varepsilon}{1+\varepsilon}}.
$$

This is a static relationship between labor supply h and the consumption-output ratio c/y in any history s^t, as a function of exogenous parameters, including the Frisch elasticity of labor supply ε, the disutility of labor supply γ, the labor share α, and the tax factor $\tau(s^t)$. It is worth stressing that this relationship holds even though productivity, government spending, and distortionary taxes may be time-varying or stochastic. Expectations of these changes are all captured by the current consumption-output ratio. For example, if productivity is currently below trend, the consumption-output ratio will be high and, to the extent that labor supply is elastic, labor supply will be low. Expectations of future tax cuts will have a similar impact on the two endogenous variables. Similarly, this model implies that if taxes are increased to fund wasteful
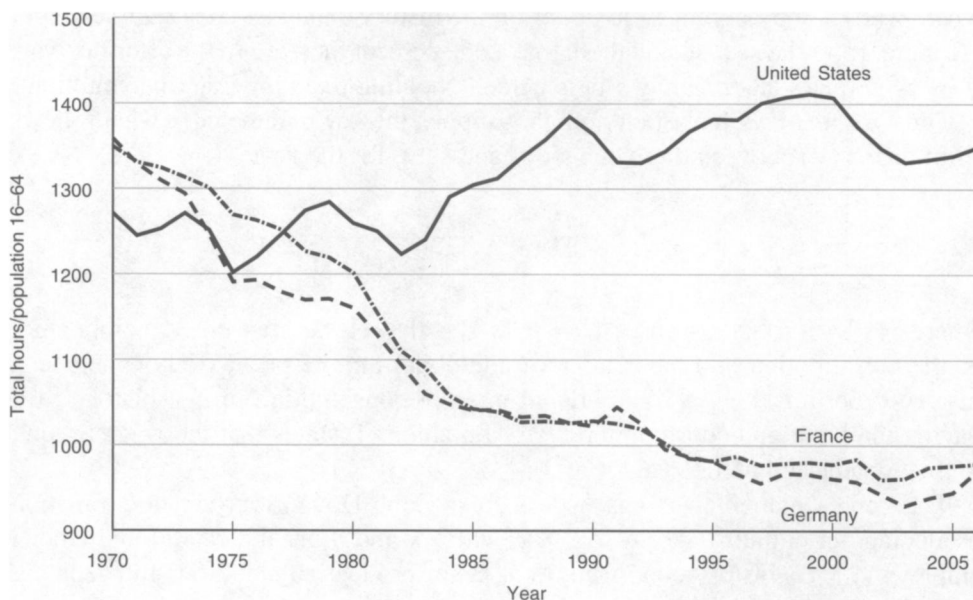
FIGURE 1

Notes: The solid line shows the ratio of total hours worked to the population aged 16–64 in the United States. The dashed line shows the ratio for Germany. The dash-dotted line shows the ratio for France. All data are from the Organization for Economic Co-operation and Development (OECD).

spending, hours need not change. Instead, the consumption-output ratio could fall by the same amount as the decline in $1 - \tau$. If the tax revenue is rebated in a lump sum to the households, the consumption-output ratio will not change, and so hours will fall. Thus, equation (7) holds regardless of whether taxes are redistributed or spent.

Prescott (2004) uses a version of equation (7) to examine the effect of tax variation on labor supply over time and across countries.[2] Figure 1 shows the number of hours worked per adult in three large economies (the United States, Germany, and France) fom 1970 to 2006. Hours worked have fallen by about 30 percent in France and Germany over this period and have increased modestly, by 7 percent, in the United States. Prescott (2004) points out that taxes rose ($1 - \tau$ fell) in Germany and France during this time period, as shown in the first column of Table 1. In addition, the consumption-output ratio increased, which may reflect expectations of future productivity growth, expectations of future tax cuts, or an increase in lump-sum transfers compared to wasteful government spending.

Equation (7) implies that, through an appropriate choice of the capital share α and the disutility of labor supply γ, I can target any desired average level for hours.

---

[2] Prescott (2004) uses a slightly different functional form for preferences, with period utility function $\log c_t + \gamma \log (100 - h_t)$, where 100 represents the available amount of time per week. He then calibrates γ to match the average number of hours worked across a broad set of countries, $\bar{h} \approx 20$. With this functional form, the Frisch elasticity of labor supply is $100/h - 1$, or about 4 on average. My choice of functional forms brings the issue of the elasticity of labor supply to the forefront of the discussion.

TABLE 1

Country	Time	Data			Theoretical h		
		$1 - \tau$	c/y	h	$\varepsilon \to 0$	$\varepsilon = 1$	$\varepsilon = 4$
Germany	1970–1974	0.48	0.66	24.6			
	1993–1996	0.41	0.74	19.3			
	log change	−0.16	0.11	−0.24	0	−0.14	−0.22
France	1970–1974	0.51	0.66	24.4			
	1993–1996	0.41	0.74	17.5			
	log change	−0.22	0.11	−0.33	0	−0.17	−0.27
US	1970–1974	0.60	0.74	23.5			
	1993–1996	0.60	0.81	25.9			
	log change	0.00	0.09	0.10	0	−0.05	−0.07

Notes: Data for $1 - \tau$, c/y, and h are from Prescott (2004, Table 2). The remaining three columns are computed using equation (7), as described in text.

During the early 1970s, the ratio $(1 - \tau)/(c/y)$ ranged from 0.73 in Germany to 0.81 in the United States, consistent with the observed similar levels of hours worked in the three countries and a common value for the technology and preference parameters. The interesting question is how the model predicts that hours should have responded to the observed change in taxes and in expectations, as summarized by changes in the consumption-output ratio.[3] Table 1 shows that in Germany, the tax factor $1 - \tau$ fell by 16 log points and the consumption-output ratio rose by 11 log points. The predicted decline in log hours is then $\varepsilon/(1 + \varepsilon)$ times the difference between the decrease in the tax factor and the decrease in the consumption-output ratio. If the Frisch elasticity of labor supply were zero, the change in τ and c/y would imply no change in hours. With $\varepsilon = 1$, it would be consistent with a 14 log point decline in hours. If $\varepsilon = 4$, hours should have declined by 22 log points. The actual decline was 24 log points. In France, the results are similar. If the Frisch elasticity of labor supply is $\varepsilon = 4$, the decline in the tax factor and the consumption-output ratio predict a 27 log point decline in hours, most of the observed 33 log point fall. In the United States, the results are more modest. The tax factor did not change, while the consumption-output ratio rose by 9 log points. This implies that hours should have fallen by 7 log points if $\varepsilon = 4$, when in fact they increased by 10 log points. Still, although the model misses the sign of the change in labor supply in the United States, it is consistent with the fact that hours changed less than in the continental European countries.

Prescott (2004) performs a similar analysis across a broader set of countries including Canada, Italy, Japan, and the United Kingdom, and with similar success. Subsequent research has examined whether Scandinavia is an outlier with high tax rates relative to the consumption-output ratio but relatively high labor supply. Kelly S. Ragan (2006) and Richard Rogerson (2007) argue that Scandinavian

---

[3] Here, I focus on hours worked per adult rather than the employment-population ratio because the cross country differences in hours worked are significantly larger. For example, according to OECD data, the number of hours worked per adult in both France and Germany was 32 log points below the level in the United States in 2006. The employment-population ratio accounted for 18 log points of the difference in France and 9 log points of the difference in Germany, with the number of hours per worker accounting for the remainder.

governments subsidize market inputs into home production and provide more trans- fers to households that supply more labor, for example, through subsidized daycare. These policies encourage households to devote time to market rather than home production. To quantify the importance of these policies, both authors develop a household production model in which time and market inputs are good substitutes in the production of services. After accounting for the uses of tax revenue, they find that the disincentive to work may not be that great in Scandinavia, bringing the model in line with data.

The critical question then is whether a Frisch labor supply elasticity of four is reasonable. The conventional answer among microeconomists is no. For example, in a prominent paper, Thomas E. MaCurdy (1981) shows that, while hours and wages are positively correlated during the life cycle, the responsiveness of hours to wages is modest. Using this source of variation, he estimates labor supply elas- ticities between 0.1 and 0.5 for white, married, prime-aged men. But this view has been attacked in recent years. Susumu Imai and Michael P. Keane (2004) argue that the measured wage is less than the shadow wage for young workers because the measured wage neglects the value of on-the-job human capital accumulation. This implies that hours are more responsive to shadow wages than to measured wages. After accounting for this, they find that the Frisch elasticity of labor sup- ply may be as high as four, particularly for older workers. Rogerson and Johanna Wallenius (2007) argue for a high elasticity of labor supply based on a different logic. They use a life-cycle model to show that indivisibilities associated with entry and exit from the labor force may make lifetime employment highly respon- sive to taxes, regardless of whether hours are responsive to wages for prime-aged workers.

In concluding this section, it is worth stressing that equation (7) implies that the responsiveness of hours to taxes is proportional to $\varepsilon/(1 + \varepsilon)$. That is, a Frisch elas- ticity of $\varepsilon = 1$ delivers half the responsiveness of hours to taxes as an infinite Frisch elasticity. To the extent that macroeconomists are convinced that the Frisch elasticity is at least equal to one, we are halfway to convergence. In any case, there is certainly broad agreement that the Frisch elasticity is important for the behavior of the labor market.

II. Cyclical Fluctuations

Until Robert E. Lucas, Jr. and Leonard A. Rapping (1969) wrote their seminal article, macroeconomists generally thought that workers supplied labor inelastically, especially in the short run. Today there is broad agreement that the same model of labor supply that works well for explaining cross-country differences is also appro- priate for analyzing the behavior of hours over the business cycle. To quantify how well the model works, simply solve equation (7) for the tax consistent with a given consumption-output ratio and level of hours:

$$(8) \qquad \tau(s^t) = 1 - \frac{\gamma}{1 - \alpha}(c(s^t)/y(s^t))h(s^t)^{\frac{1+\varepsilon}{\varepsilon}}.$$
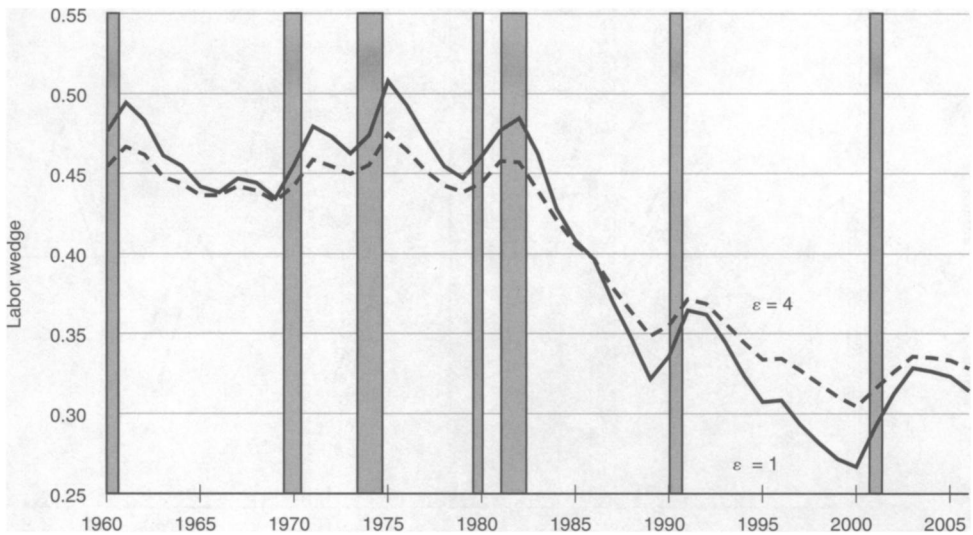
FIGURE 2

Notes: The US labor wedge from equation (8). The solid line shows $\varepsilon = 1$, and the dashed line shows $\varepsilon = 4$. In both cases, I fix the remaining parameters to ensure that the average labor wedge is 0.40. The gray bands show National Bureau of Economic Research (NBER) recession dates.

The key observation is that hours and the consumption-output ratio can be measured at high frequencies, and so by making appropriate assumptions about γ, α, and ε, one can back out the implicit wedge $\tau(s^t)$. This is the labor wedge. The idea of using a version of equation (8) to measure the wedge between the marginal rate of substitution and the marginal product of labor is not new. A diverse group of authors have converged on this approach, including Michael Parkin (1988), Julio J. Rotemberg and Michael Woodford (1991) and (1999), Robert E. Hall (1997), Casey B. Mulligan (2002), and V. V. Chari, Patrick J. Kehoe, and Ellen R. McGrattan (2007).

To implement it, I use an OECD measure of hours that accounts for production and nonproduction workers, as well as the self-employed.[4] I use the ratio of nominal personal consumption expenditures to nominal GDP from the National Income and Product Accounts to measure the consumption-output ratio. For a variety of values of ε, I fix $\gamma/(1 - \alpha)$ so as to ensure that the average labor wedge is 0.40 during the period when data are available, 1960 to 2006 for the United States, consistent with the numbers in Table 1. Figure 2 shows the results, while Figure 3 shows the detrended wedge.

At low frequencies, Figure 2 shows a substantial reduction in the labor wedge during the 1980s. Arguably, this was associated with the Reagan tax reforms. More pertinent, Figure 3 shows a sharp increase in the labor wedge during every recession, as indicated by the gray bars. The magnitude of the implied cycles in the labor

---

[4] The data for hours are the product of the employment-population ratio, using labor force status by sex and age, and average annual hours actually worked per worker. They are available from http://stats.oecd.org/wbos/default.aspx.
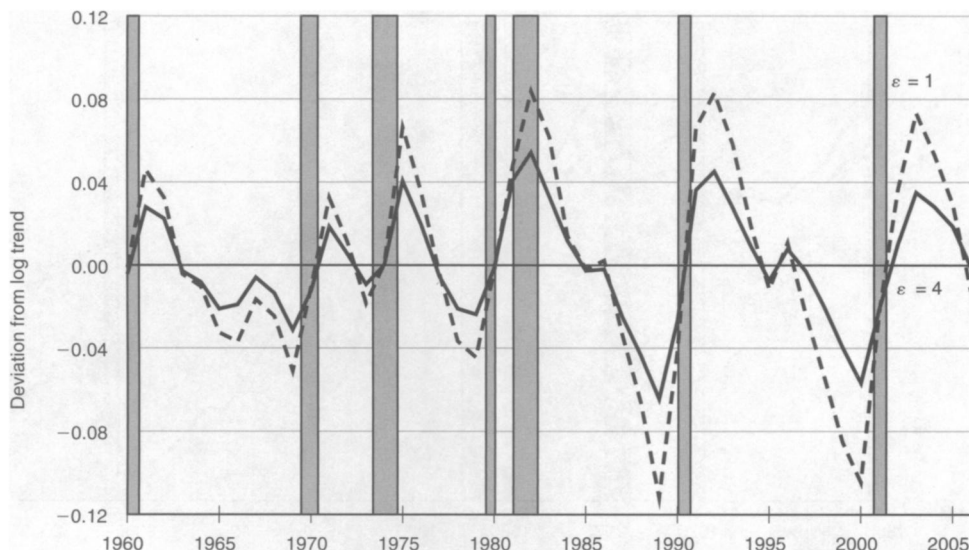
FIGURE 3

Notes: Deviation of the labor wedge from log trend, HP filter with parameter 100. The solid line shows $\varepsilon = 1$, and the dashed line shows $\varepsilon = 4$. In both cases, I fix the remaining parameters to ensure that the average labor wedge is 0.40. The gray bands show NBER recession dates.

wedge depends on the elasticity of labor supply. For example, with $\varepsilon = 1$, the 1990 recession is associated with a 20 percent jump in the labor wedge relative to trend, while with $\varepsilon = 4$, the increase is about half as large.

There are a number of possible explanations for this pattern. The most obvious is that labor and consumption taxes rise in recessions. This hypothesis is not a priori unreasonable. In their comparison of the cyclical behavior of hours predicted by Harald Uhlig (2003) on the one hand and by Kaiji Chen, Ayşe İmrohoroğlu, and Selahattin İmrohoroğlu (2007) on the other, Ellen R. McGrattan and Prescott (2007) note that the latter paper fits the data much better than the former, and the main difference between the two approaches is the inclusion of variation in taxes. Using a different methodology based on the Christina D. Romer and David H. Romer (2007) narrative analysis of tax policy, Karel Mertens and Morten O. Ravn (2008) conclude that tax shocks account for 18 percent of the variance of output at business cycle frequencies. Perhaps most provocatively, they conclude that the 1982 recession was caused by workers' anticipation of future tax cuts. But despite these recent papers, most economists seem to be skeptical that tax movements alone can explain the observed variation in the labor wedge.

The second possibility is that either the MRS or MPL is misspecified. The specification of the MPL depends only on the assumption of a Cobb-Douglas aggregate production function. Macroeconomists are justifiably reluctant to abandon that assumption because it ensures that the capital and labor shares of national income as well as the interest rate are constant, consistent with the Nicholas Kaldor (1957) growth facts.

The specification of household preferences is also tightly constrained by long-run restrictions. For simplicity, maintain the assumption that preferences are time-separable and represent the period utility function as $u(c,h)$. If they are additionally separable between consumption and leisure, balanced growth—the absence of a long-run trend in hours—requires $u(c,h) = \log c - v(h)$. The balanced growth restriction seems in line with the trends in Figure 1, while my specification of $v(h) = \gamma h^{(1+\varepsilon)/\varepsilon}$ ensures a constant Frisch labor supply elasticity. Since the labor supply elasticity does not affect the behavior of the labor wedge (Figure 3) substantially, this restriction seems innocuous.

Instead, I relax the assumption of additive separability between consumption and leisure. To be consistent with balanced growth and a constant Frisch elasticity ε, the period utility function must satisfy

$$u(c,h) = \frac{c^{1-\sigma}\left(1 + (\sigma-1)\left(\gamma\varepsilon/(1+\varepsilon)\right)h^{(1+\varepsilon)/\varepsilon}\right)^\sigma - 1}{1-\sigma},$$

with $\sigma > 0$ denoting the coefficient of relative risk aversion and $\gamma > 0$ denoting the disutility of labor supply. These parameter restrictions ensure that utility is increasing and concave in consumption and decreasing and concave in hours of work. The limit as $\sigma \to 1$ nests the time separable case in equation (1). The case where $\sigma > 1$ is of particular interest, since this implies the marginal utility of consumption is higher when households work more, as predicted by standard models of time allocation (Becker 1965). In any case, with this functional form for preferences, the labor wedge satisfies

$$(9) \qquad \tau(s^t) = 1 - \frac{\gamma\sigma}{1-\alpha} \frac{(c(s^t)/y(s^t))h(s^t)^{(1+\varepsilon)/\varepsilon}}{1 + (\sigma-1)\left(\gamma\varepsilon/(1+\varepsilon)\right)h(s^t)^{(1+\varepsilon)/\varepsilon}},$$

a modest generalization of equation (8). To understand the quantitative implications of this expression, fix the labor share at the conventional value of $1 - \alpha = 2/3$. For different values of σ and ε, choose the disutility of work parameter γ to ensure an average labor wedge of 0.40 in the United States since 1960. Figure 4 shows the time series behavior of the labor wedge with the Frisch elasticity fixed at one. The solid line corresponds to the limit as σ converges to one, the additively separable case that I analyzed before, while the dashed line shows $\sigma = 4$. Raising risk aversion modestly reduces the magnitude of fluctuations in the labor wedge but does not qualitatively change the results.[5] I do not show the results with a higher elasticity of labor supply, but they are similar.

Additionally, the microeconomic behavior of the model is unreasonable when σ is much larger than one. Consider the following thought experiment: a worker who normally supplies 2,000 hours of labor per year anticipates next year she will not be able work. With complete markets, she wishes to keep the marginal utility

---

[5] Additive separability is also not very important for the cross-country results in Table 1. With a Frisch elasticity of 1, the predicted decline in hours in Germany is 15 log points when $\sigma = 4$ (compared with 14 log points with $\sigma \to 1$). In France, it is 18 log points (compared with 17). In the United States, it remains at 5 log points.
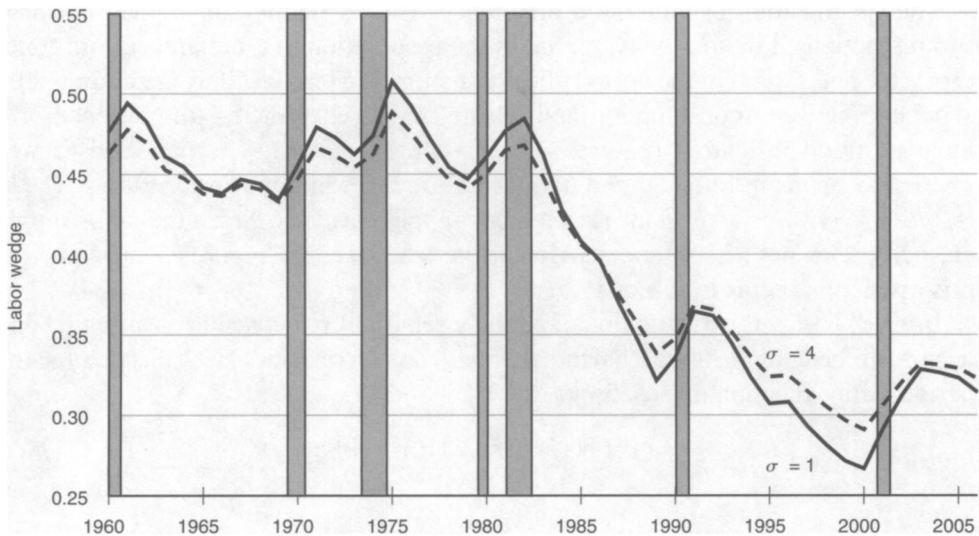
FIGURE 4

Notes: The US labor wedge from equation (9). The solid line shows $\sigma \to 1$ and the dashed line shows $\sigma = 4$. The Frisch elasticity is $\varepsilon = 1$, the labor share is $1 - \alpha = 2/3$, and I adjust the disutility of work γ to ensure that the average labor wedge is 0.40. The gray bands show NBER recession dates.

of consumption constant through this episode. How much should her consumption decline when she stops working? With $\sigma = 1$, additive separability, consumption remains constant. With $\varepsilon = 1$ and $\sigma = 1.3$, consumption falls by 16 log points. At $\sigma = 2$, it falls by 35 log points; and at $\sigma = 4$, it falls by 53 log points. This can be compared to the drop in consumption expenditures at retirement. Mark Aguiar and Erik Hurst (2005) find that food consumption expenditures drop by about 17 percent, accompanied by a 53 percent increase in the time spent on food production. This type of evidence severely restricts the curvature in the utility function and thus the fruitfulness of alternative specifications of the MRS.

The third possible explanation for the behavior of the labor wedge is that the disutility of work, γ, is time varying. Hours are low relative to the consumption-output ratio during recessions because the disutility of work is high. Like many economists, I have a strong prior belief that this is a poor explanation for the pattern in Figure 3. Although households may differ in their disutility of work and the disutility may change over time for some households, one would expect those movements to average out in a large economy. Mine is not a novel view. N. Gregory Mankiw (1989, footnote 1) echoed Franco Modigliani (1977) in writing, "Alternatively, one could explain the observed pattern by positing that tastes for consumption relative to leisure vary over time. Recessions are then periods of 'chronic laziness.' As far as I know, no one has seriously proposed this explanation of the business cycle." Mankiw wrote prematurely. Rotemberg and Woodford (1997) give an important role to an unobserved demand shock in explaining aggregate fluctuations, a combination of a preference shock and a shock to government spending. Christopher J. Erceg, Dale W. Henderson, and Andrew T. Levin (2000) and Frank Smets and Raf Wouters (2003)

also have quantitatively important preference shocks in their models of monetary policy. More recently, Jordi Galí and Pau Rabanal (2004) find, in their estimated dynamic stochastic general equilibrium model, that a preference shock explains 57 percent of the variance of output and 70 percent of the variance of hours. While Figure 3 explains why they need this shock to make their model fit the data, it does not make the resolution intellectually appealing.

A closely related possibility is that workers do not take the wage as given when deciding how much to work, but instead have time-varying market power in the labor market. In a common formulation, household $i \in [0, 1]$ is endowed with a heterogeneous type of labor and supplies $\eta_i(s^t)$ units of it in state s^t by setting the wage $\omega_i(s^t)$. A representative price- and wage-taking firm produces a homogeneous intermediate input h using a technology with a history-contingent elasticity of substitution $\theta(s^t) > 1$ between the heterogeneous types of labor,

$$(10) \qquad h(s^t) = \left(\int_0^1 \eta_i(s^t)(\theta(s^t) - 1)/\theta(s^t) \, di \right)^{\theta(s^t)/\theta(s^t) - 1}.$$

Let $w(s^t)$ denote the rental price of the intermediate good. Then the intermediate goods producer chooses $\eta_i(s^t)$ for each type of labor i to maximize $w(s^t)h(s^t) - \int_0^1 \omega_i(s^t)\eta_i(s^t)$. The solution to this profit maximization problem gives the inverse demand curve for each type of labor,

$$(11) \qquad \omega_i(s^t) = w(s^t)(h(s^t)/\eta_i(s^t))^{1/\theta(s^t)}.$$

Households choose $\eta_i(s^t)$ optimally, earning pre-tax labor income $\eta_i(s^t)\omega_i(s^t)$ in history s^t, where $\omega_i(s^t)$ solves equation (11). Replacing this in the budget constraint equation (2) and solving the household's problem delivers a new first-order condition for labor supply,

$$\frac{\gamma\theta(s^t)}{\theta(s^t) - 1} \eta_i(s^t)^{1/\varepsilon} = \lambda(s^t)(1 - \tau_h(s^t))w(s^t).$$

Combining this with the first-order condition for consumption in equation (3), I find that the price of the intermediate input is a history-contingent markup over the marginal rate of substitution between consumption and leisure:

$$(12) \qquad w(s^t) = \frac{\gamma\theta(s^t)c(s^t)\eta_i(s^t)^{1/\varepsilon}}{(\theta(s^t) - 1)(1 - \tau(s^t))}.$$

Given the symmetry of the problem, all households choose the same labor supply, $\eta_i(s^t) = h(s^t)$. To close the model, assume final goods producers combine $k(s^t)$ units of capital and $n(s^t)$ units of the intermediate input using a Cobb-Douglas production function to produce final output. The profit function is unchanged from equation (5) and so equation (6) remains the first-order condition for the use of the intermediate

input. Combine this with equation (12) and the intermediate goods market-clearing condition $n(s^t) = h(s^t)$ to obtain a generalization of equation (8):

$$(13) \qquad \tau(s^t) = 1 - \frac{\gamma\theta(s^t)}{(1-\alpha)(\theta(s^t)-1)}(c(s^t)/y(s^t))h(s^t)^{(1+\varepsilon)/\varepsilon}.$$

If one treats the elasticity of substitution $\theta(s^t)$ as a residual, any path for the consumption-output ratio and hours is consistent with a constant labor wedge $\tau(s^t)$. That is, shocks to $\theta(s^t)$ and hence to workers' monopoly power in the labor market can "explain" the pattern in Figure 3. According to this view, recessions are periods when the elasticity of substitution between different varieties of labor is unusually low. Wage markups rise, increasing the wedge between the marginal rate of substitution and the marginal product of labor. Of course, equation (13) shows that this is observationally equivalent to an increase in the disutility of leisure γ. The idea that recessions are periods of widespread monopolization of the labor market is empirically as implausible as the idea that recessions are periods of chronic laziness. Despite this, a number of recent papers have emphasized this as an important source of business cycle shocks including Smets and Wouters (2003) and (2007) and Galí, Mark Gertler, and David López-Salido (2007). In Smets and Wouters (2007), the wage markup shock accounts for 20 percent of the variance in output and over half the variance in inflation at a 10-quarter horizon.

There are two remaining explanations for movements in the labor wedge. First, it is conceivable that the assumption of a representative household and a representative firm neglects an important role for microeconomic heterogeneity. Although it is clear that the representative agent approach misses much of the richness that we observe in the world, it is less clear that this has important effects on the business cycle properties of models. Notably, Per Krusell and Anthony A. Smith, Jr. (1998) develop a nonrepresentative agent business cycle model with uninsurable idiosyncratic risk and discount factor heterogeneity. Their main finding is that the business cycle properties of the model are virtually identical to an analogous representative agent model. Although it is possible that some other nonrepresentative agent model will deliver a countercyclical labor wedge, this approach currently does not seem promising.

Second, the MRS and MPL may not be equal to each other either because the MPL is not equal to the wage, the MRS is not equal to the wage, or both. That is, the labor market is not competitive. For example, one might hope that sticky wages would deliver an endogenous labor wedge. Consider a variant of the intermediate goods production function in equation (10), but with a constant elasticity of substitution θ, and suppose that wages are sticky, e.g. at least some nominal or real wages are determined before the current shock realization. The labor market clears in history s^t via each household i's commitment to supply whatever labor $\eta_i(s^t)$ intermediate goods producers demand at the predetermined wage $\omega_i(s^t)$, i.e. from equation (11). This breaks the short-run link between the wage and the MRS. Following a negative productivity shock, workers are unable to provide as much labor as they would like to, raising the measured labor wedge.

There are at least two problems with this theory of the labor wedge. First, sticky wages imply that workers and firms do not exploit all of the gains from trade, a
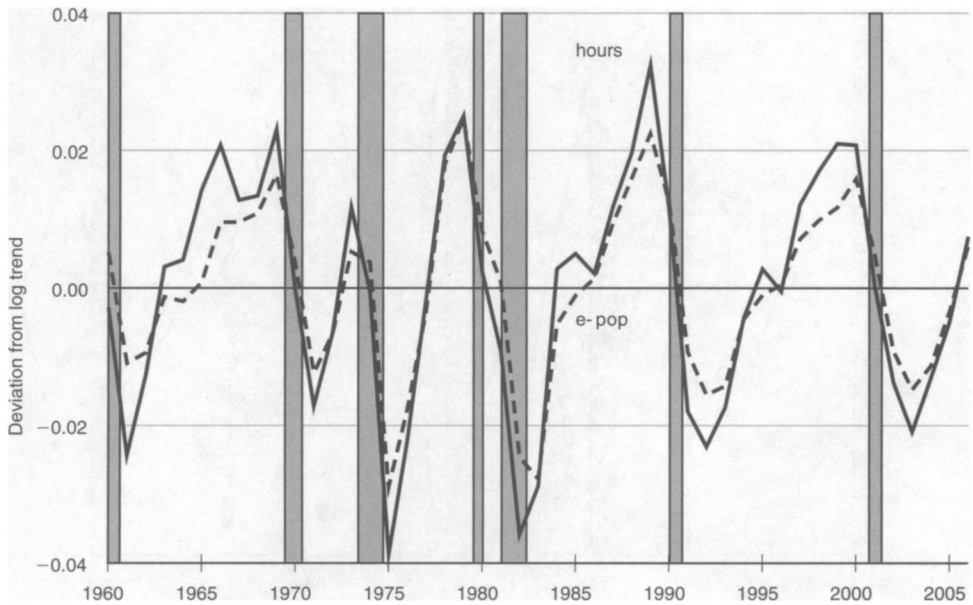
FIGURE 5

Notes: Deviation of the hours and the e-pop ratios from log trend, HP filter with parameter 100. The solid line shows the deviation of hours from log trend, and the dashed line shows the deviation of the e-pop ratio. The gray bands show NBER recession dates.

version of the Robert J. Barro (1977) critique. Indeed, because most workers are in long-term employment relationships, it is unclear why their employers would choose to determine employment from the static labor demand schedule in each history. There may be significant gains to explicit or implicit long-term contracts. Second, the explanation seems to be a quantitative failure. The important role played by exogenous shocks to the disutility of labor or workers' monopoly power in the sticky wage models discussed above (Rotemberg and Woodford 1997; Erceg, Henderson, and Levin 2000; Smets and Wouters 2003, 2007; Galí and Rabanal, 2004; Galí, Gertler, and López-Salido 2007) indicates that the endogenous mechanism alone cannot generate the magnitude of the observed labor wedge.

Search and matching models based on Christopher A. Pissarides (1985) and Dale T. Mortensen and Pissarides (1994), in which wages are determined in decentralized meetings, provide an alternative explanation for why the MRS and MPL are unequal and are an ideal laboratory for exploring cyclical fluctuations in labor market outcomes. To understand why I say that this is an ideal laboratory, note first that search costs introduce a nonconvexity into households' decision problem which focuses attention on the binary decision of whether to work rather than the continuous decision of how many hours to work each week. In contrast to the cross-country evidence that I discuss in footnote 3, the data indicate that this focus is appropriate. Figure 5 shows that when hours are 1 percent above trend, the employment-population (e-pop) ratio is also nearly 1 percent above trend. Thus, most business cycle frequency fluctuations in hours are accounted for by fluctuations in the e-pop ratio, rather than
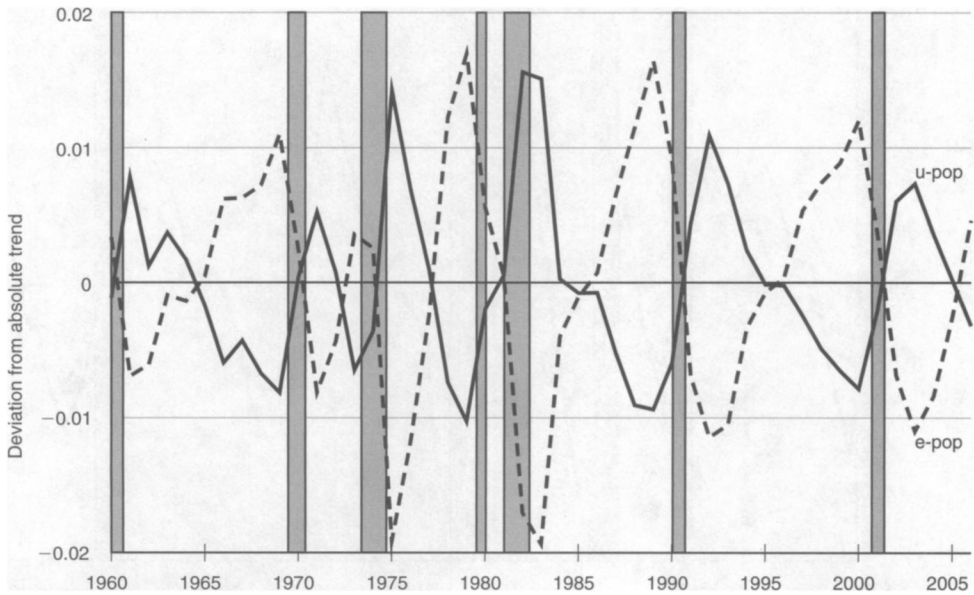
FIGURE 6

Notes: Deviation of the e-pop and u-pop ratios from trend, HP filter with parameter 100. The solid line shows the deviation of the u-pop ratio from trend, and the dashed line shows the deviation of the e-pop ratio. The gray bands show NBER recession dates.

fluctuations in the number of hours per employee. Less crucially, search models also often focus on the margin between employment and unemployment, neglecting entry and exit from the labor force. Again, this is empirically reasonable at business cycle frequencies. Figure 6 shows that when the e-pop ratio is 1 percentage point above trend, the unemployment-population (u-pop) ratio is approximately 1 percentage point below trend. The third category, nonparticipation in the labor market, is comparatively acyclic. Most business cycle frequency fluctuations in e-pop ratio are offset by equal movements in the u-pop ratio, so movements in and out of the labor force are comparatively unimportant at business cycle frequencies.

An implication of the binary decision about whether to work is that the Frisch elasticity of labor supply is effectively infinite (Gary D. Hansen, 1985; Rogerson, 1988). Assume for simplicity that households are made up of a unit measure of individuals, each with preferences given by equation (1). Also assume that labor is indivisible, so $h(s^t) \in \{0, 1\}$ for each member of the household. Then if household members pool their income to insure each other against shocks to their labor income, the household acts as if it has preferences

$$U(\{c, e\}) = \sum_{t=0}^{\infty} \beta^t \left(\sum_{s^t} \pi(s^t) (\log c(s^t) - \tilde{\gamma} e(s^t)) \right),$$

where $e(s^t)$ is the fraction of household members who are employed in history s^t, and $\tilde{\gamma} \equiv \gamma \varepsilon / (1 + \varepsilon)$ measures the disutility of working. Since the household's preferences

are linear in its employment rate, it orders consumption and leisure choices exactly like a household with divisible labor and an infinite elasticity of labor supply. Still, even an infinite elasticity of labor supply does not eliminate the labor wedge.

Instead, the important feature of matching models is that search frictions create a bilateral monopoly situation between workers and firms. In a standard formulation, workers and firms engage in a time consuming search for partners before negotiating a wage. Once they have sunk this cost, there is a range of wages at which both prefer to match rather than breakup. Loosely speaking, any wage in this range is bigger than the MRS but smaller than the MPL.

A critical question is how wages are determined. A standard assumption is that the worker and firm bargain over the gains from trade, splitting the surplus according to the Nash bargaining solution. Using a different specification for preferences and a different metric for model evaluation, Shimer (2005) finds that a calibrated version of the Pissarides (1985) matching model generates only very small fluctuations in labor market outcomes in response to plausible productivity shocks. More closely related to my analysis in this essay, Olivier J. Blanchard and Galí prove that, with the preferences in equation (1), productivity shocks affect neither the labor wedge nor the unemployment rate. Instead, the wage, the MRS, and the MPL all move in proportion to the underlying shock. This implies that search frictions, per se, cannot explain the pattern in Figure 3.

But other wage setting procedures are no less plausible than the Nash bargaining solution and have vastly different implications for the behavior of the model. In a framework similar to Shimer (2005), Hall (2005) shows that if wages are rigid because of a social norm, unemployment is extremely sensitive to underlying shocks. He stresses that this type of wage rigidity is not susceptible to the Barro (1977) critique. Blanchard and Galí (2006) consider a real wage rigidity that makes the wage move less than one-for-one with the shock. Firms respond to relatively low wages during booms by creating many new jobs, driving down the unemployment rate. However, this also implies that part of the productivity increase is spent on additional job creation. Consumption then increases by less than productivity, generating a countercyclical labor wedge. Gertler and Antonella Trigari (2006) reach a similar conclusion in a model with overlapping wage contracts that are not contingent on the path of productivity shocks.

This is a new and rapidly changing research area, so whether this is a satisfactory explanation for fluctuations in the labor wedge remains an open question. My view is that macroeconomic data on consumption, output, hours, and wages will not provide a conclusive answer. While the microeconomic evidence is sparse, Pissarides (2007) reviews the relevant literature and offers a skeptical appraisal of the evidence for important real wage rigidities. If he is right, macroeconomists may need to look beyond search models for an explanation of the labor wedge.

III. Conclusion

This essay advocates focusing on a particular aspect of macroeconomic models, the labor wedge. The advantage to this approach is that the behavior of the labor wedge depends only on a few details of the model—households maximize utility,

firms maximize profits, and markets are competitive and clear—sidestepping the need to specify the nature of shocks and the formation of expectations. To the extent that macroeconomists accept the assumptions that households and firms optimize, one can focus attention on the third assumption—do nonmarket clearing models, such as search models, provide a compelling explanation for cyclical patterns in the labor wedge? Answering this question is likely to remain an important research topic in the coming years.

REFERENCES

Aguiar, Mark, and Erik Hurst. 2005. "Consumption versus Expenditure." *Journal of Political Economy*, 113(5): 919–48.

Barro, Robert J. 1977. "Long-Term Contracting, Sticky Prices, and Monetary Policy." *Journal of Monetary Economics*, 3(3): 305–16.

Becker, Gary S., 1965. "A Theory of the Allocation of Time." *The Economic Journal*, 75(299): 493–517.

Blanchard, Olivier J., and Jordi Galí. 2006. "Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment." MIT Department of Economics Working Paper 06-22.

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

Chen, Kaiji, Ayşe İmrohoroğlu, and Selahattin İmrohoroğlu, 2007. "A Quantitative Assessment of the Decline in the U.S. Savings Rate." Unpublished.

Erceg, Christopher J., Dale W. Henderson, and Andrew T. Levin. 2000. "Optimal Monetary Policy with Staggered Wage and Price Contracts." *Journal of Monetary Economics*, 46(2): 281–313.

Galí, Jordi, Mark Gertler, and J. David Lopez-Salido. 2007. "Markups, Gaps, and the Welfare Costs of Business Fluctuations." *Review of Economics and Statistics*, 89(1): 44–59.

Galí, Jordi, and Pau Rabanal, 2004. "Technology Shocks and Aggregate Fluctuations: How Well Does the RBC Model Fit Postwar US Data?" In *NBER Macroeconomics Annual*. Vol. 19, ed. Mark Gertler and Kenneth Rogoff, 225–318. Cambridge, MA: MIT Press.

Gertler, Mark, and Antonella Trigari. 2006. "Unemployment Fluctuations with Staggered Nash Wage Bargaining." National Bureau of Economic Research Working Paper 12498.

Hall, Robert E. 1997. "Macroeconomic Fluctuations and the Allocation of Time." *Journal of Labor Economics*, 15(1): S223–50.

Hall, Robert E. 2005. "Employment Fluctuations with Equilibrium Wage Stickiness." *American Economic Review*, 95(1): 50–65.

Hansen, Gary D. 1985. "Indivisible Labor and the Business Cycle." *Journal of Monetary Economics*, 16(3): 309–27.

Imai, Susumu, and Michael P. Keane. 2004. "Intertemporal Labor Supply and Human Capital Accumulation." *International Economic Review*, 45(2): 601–41.

Kaldor, Nicholas. 1957. "A Model of Economic Growth." *The Economic Journal*, 67(268): 591–624.

Krusell, Per, and Anthony A. Smith, Jr. 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 106(5): 867–96.

Lucas, Robert E., Jr., and Leonard A. Rapping. 1969. "Real Wages, Employment, and Inflation." *Journal of Political Economy*, 77(5): 721–54.

MaCurdy, Thomas E. 1981. "An Empirical Model of Labor Supply in a Life-Cycle Setting." *Journal of Political Economy*, 89(6): 1059–85.

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

McGrattan, Ellen R., and Edward C. Prescott, 2007. "Technical Appendix: Unmeasured Investment and the Puzzling US Boom in the 1990s." Federal Reserve Bank of Minneapolis Staff Report 395.

Mertens, Karel, and Morten O. Ravn. 2008. "The Aggregate Effects of Anticipated and Unanticipated U.S. Tax Policy Shocks: Theory and Empirical Evidence." Unpublished.

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

Mortensen, Dale T., and Christopher A. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies*, 61(3): 397–415.

Mulligan, Casey B. 2002. "A Century of Labor-Leisure Distortions." National Bureau of Economic Research Working Paper 8774.

Parkin, Michael. 1988. "A Method for Determining Whether Parameters in Aggregative Models Are Structural." *Carnegie-Rochester Conference Series on Public Policy*, 29: 215–52.

Pissarides, Christopher A. 1985. "Short-Run Equilibrium Dynamics of Unemployment Vacancies, and Real Wages." *American Economic Review*, 75(4): 676–90.

Pissarides, Christopher A. 2007. "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?" Unpublished.

Prescott, Edward C. 2004. "Why Do Americans Work So Much More Than Europeans?" *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(1): 2–13.

Ragan, Kelly S. 2006. "Taxes, Transfers, and Time Use: Fiscal Policy in a Household Production Model." PhD diss. University of Chicago.

Rogerson, Richard. 1988. "Indivisible Labor, Lotteries and Equilibrium." *Journal of Monetary Economics*, 21(1): 3–16.

Rogerson, Richard. 2007. "Taxation and Market Work: Is Scandinavia an Outlier?" *Economic Theory*, 32(1): 59–85.

Rogerson, Richard, and Johanna Wallenius. 2007. "Micro and Macro Elasticities in a Life Cycle Model with Taxes." National Bureau of Economic Research Working Paper 13017.

Romer, Christina D., and David H. Romer. 2007. "A Narrative Analysis of Postwar Tax Changes." Unpublished.

Rotemberg, Julio J., and Michael Woodford. 1991. "Markups and the Business Cycle." In *NBER Macroeconomics Annual*, ed. Olivier J. Blanchard and Stanley Fischer, 63–129. Cambridge, MA: MIT Press.

Rotemberg, Julio J., and Michael Woodford. 1997. "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy." In *NBER Macroeconomics Annual*. Vol. 12, ed. Ben Bernanke and Julio J. Rotemberg. 297–346. Cambridge, MA: MIT Press.

Rotemberg, Julio J., and Michael Woodford. 1999. "The Cyclical Behavior of Prices and Costs." In *Handbook of Macroeconomics*. Vol. 1B, ed. John B. Taylor and Michael Woodford. Cambridge, MA: MIT Press: 1051–1135.

Shimer, Robert. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *American Economic Review*, 95(1): 25–49.

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): 1123–75.

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

Uhlig, Harald, 2003. "How Well do we Understand Business Cycles and Growth? Examining the Data with a Real Business Cycle Model." In *Empirische Wirtschaftsforschung: Methoden und Anwendungen*, Vol. 32 of *Wirtschaftswissenschaftliches Seminar Ottobeuren*. ed. Wolfgang Franz, Hans Jürgen Ramser and Manfred Stadler, 295–319.