MARKUPS, GAPS, AND THE WELFARE COSTS OF BUSINESS FLUCTUATIONS

Jordi Galí, Mark Gertler, and J. David López-Salido*

Abstract—We present a simple theory-based measure of the variations in aggregate economic efficiency: the gap between the marginal product of labor and the household's consumption leisure tradeoff. We show that this indicator corresponds to the reciprocal of the markup of price over social marginal cost, and give some evidence in support of this interpretation. We then show that, with some auxiliary assumptions, our gap variable may be used to measure the efficiency costs of business fluctuations. We find that the latter costs are modest on average. However, to the extent that the flexible price equilibrium is distorted, the gross efficiency losses from recessions and gains from booms may be large. Indeed, we find that the major recessions involved large efficiency losses. These results hold for reasonable parameterizations of the Frisch elasticity of labor supply, the relative risk aversion, and steady-state distortions.

I. Introduction

To the extent that there exist price and wage rigidities, or possibly other types of market frictions, the business cycle is likely to involve inefficient fluctuations in the allocation of resources. Specifically, the economy may oscillate between expansionary periods when the volume of economic activity is close to the social optimum, and recessions that feature a significant drop in production relative to the first best. In this paper we explore this hypothesis by developing a simple measure of aggregate inefficiency and examining its cyclical properties. The measure we propose—which we call the *inefficiency gap,* or the *gap,* for short—is based on the size of the wedge between the marginal product of labor and the marginal rate of substitution between consumption and leisure. Deviations of this gap from 0 reflect an inefficient allocation of employment. By constructing a time series measure of the inefficiency gap, we are able to obtain some insight into both the nature and the welfare costs of business cycles.

From a somewhat different perspective, we show that the inefficiency gap corresponds to the reciprocal of the markup of price over social marginal cost. Procyclical movements in the inefficiency gap accordingly mirror countercyclical movements in this markup. Our approach, however, differs from much of the recent literature on business cycles and markups by using the household's marginal rate of substitution between consumption and leisure to measure the price of labor, as opposed to wages.[1] As a matter of theory, of course, the household's consumption leisure tradeoff is the appropriate measure of the true social cost of labor. Wage data are not appropriate if either wages are not allocative or labor market frictions are present that drive a wedge between market wages and the labor supply curve. As we demonstrate, our markup construct is highly countercyclical. In addition, it also leads directly to a measure of aggregate efficiency costs at each point in time.

Our approach builds on a stimulating paper by Hall (1997) that analyzes the cyclical behavior of the neoclassical labor market equilibrium. Specifically, Hall first demonstrates that the business cycle is associated with highly procyclical movements in the difference between the observable component of the household's marginal rate of substitution and the marginal product of labor. He then presents some evidence to suggest that this difference is of central importance to employment fluctuations. Also relevant is Mulligan (2002), who examines essentially the same measure of the labor market wedge, though focusing on its low-frequency movements. Specifically, he constructs an annual series of this variable, using data spanning more than a century. He finds that marginal tax rates correlate well at low frequencies with this labor market wedge. Finally, Chari, Kehoe, and McGrattan (2004) find that the labor market wedge plays a critical role in accounting for the drop in employment during the Great Depression.

As with Hall, we focus on the behavior of the labor market wedge at the business cycle frequency. We differ in several important ways, however. First, his framework treats this wedge simply as an exogenous driving force, interpretable, for example, as reflecting shifts in preferences.[2] We instead stress countercyclical markup variations as the key factor accounting for the cyclical fluctuations in this variable and present evidence in support of this general hypothesis. Second, given our *markup interpretation,* we are able to use the Hall residual as the basis for a measure of the efficiency costs of business cycles.

In particular, with some auxiliary assumptions, it is possible to derive a measure of the lost surplus in the labor market at each point in time, based directly on movements

Received for publication October 29, 2003. Revision accepted for publication November 1, 2005.

* CREI and Universitat Pompeu Fabra, New York University, and Federal Reserve Board and CEPR, respectively.

We thank Jeff Amato, Susanto Basu, Bob Hall, Andy Levin, Casey Mulligan, Jonathan Parker, Julio Rotemberg, and several anonymous referees for helpful comments, as well as participants in seminars at LSE, Bocconi, IIES, UPF, Bank of Spain, Bank of England, NBER Summer Institute, CERGE, EUI, CEU, UAB, and the ECB Workshop on DSGE Models. Galí is thankful to CREA-Barcelona Economics, the Bank of Spain, the MCYT, and DURSI for financial support. Gertler acknowledges the support of the NSF and the C. V. Starr Center. The opinions expressed here are solely those of the authors and do not necessarily reflect the views of the Board of Governors of the Federal Reserve System or of anyone else associated with the Federal Reserve System.

[1] See Rotemberg and Woodford (1999) for a survey of the literature on business cycles and countercyclical markups.

[2] To organize his approach, Hall (1997) modeled the labor market residual as an unobserved preference shock, though he did not take this hypothesis literally, but rather as a starting point for subsequent analysis. There has been a tendency in subsequent literature, however, (for example, Holland & Scott, 1998; Francis & Ramey, 2001; Uhlig, 2002) to interpret this residual as an exogenous preference shock. Earlier literature as well offered a similar interpretation (for example, Baxter & King, 1991). Our analysis will suggest that this residual cannot simply reflect exogenous preference shifts.

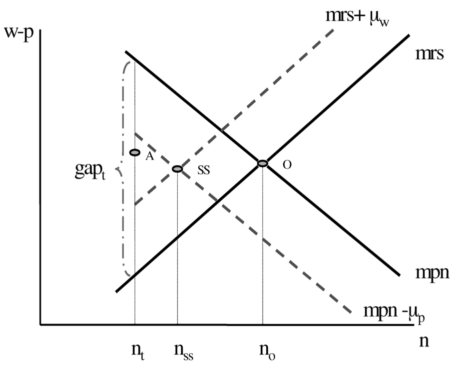
in our gap variable. Fluctuations generate efficiency costs on average because, as we show, the surplus lost from a decline in employment below its natural level exceeds the gain from a symmetric rise above its natural level. In this respect, our approach differs significantly from that of Lucas (1987, 2003), who examines the welfare costs of consumption variability associated with the cycle. Whereas the Lucas measure does not really take account of the sources of fluctuations, our measure instead isolates the costs associated with the inefficient component of fluctuations. Accordingly, our metric may give a better sense of the potential gains from improved stabilization policy.

A significant additional feature is that our approach permits not only a measure of the costs of fluctuations on average, but also an assessment of the costs of particular episodes. We find, for example, that although the efficiency costs of fluctuations are not large when averaged across booms and recessions, the gross gains from booms and losses from recessions can indeed be quite large. Indeed, as we show, our methodology suggests that the U.S. economy experienced large efficiency costs during both the 1974–1975 and the 1980–1982 recession. This consideration is highly relevant because it may be that the main benefit from good stabilization policies may be to avoid severe recessions. To the extent that central banks have had either good skill or good luck in keeping to a minimum the number of severe downturns, it may be that on average the costs of fluctuations are not large. This kind of unconditional calculation, however, masks the kind of losses that can emerge if luck and/or skill suddenly turn bad. For this reason, an examination of episodes where matters clearly did seem to go awry can shed light on the importance of good policy management.

In section II we develop a framework for measuring the inefficiency gap in terms of observables, conditional on some reasonably conventional assumptions about preferences and technology. We also show that it is possible to decompose the gap into price and wage markup components. In section 3 we present empirical measures of this variable for the postwar U.S. economy. The inefficiency gap exhibits large procyclical swings. In addition, under the assumption that wages are allocational, most of its variation is associated with countercyclical movements in the wage markup.[3] The price markup shows, at best, a weak contemporaneous correlation. Under some alternatives to our baseline case, the price markup does move countercyclically. However, movements in the wage markup still dominate the overall movements in the gap.

In section IV we consider the possibility that purely exogenous factors (such as unobserved preference shifts)

FIGURE 1.—THE GAP: A DIAGRAMMATIC EXPOSITION



underlie the variation in our gap measures. Specifically, we present evidence that suggests that our gap variable is endogenous and thus cannot simply reflect exogenous variation in preferences. The evidence is instead consistent with our maintained hypothesis that endogenous variation in markups is largely responsible for the movement in the inefficiency gap. In section 5 we then use this link to examine both the unconditional efficiency costs of recessions and the conditional costs associated with the major boom-bust episodes. Concluding remarks are in section VI.

II. The Gap and Its Components: Theory

Let the *inefficiency gap* (henceforth, the *gap*) be defined as follows:

$$gap_t = mrs_t - mpn_t, \tag{1}$$

where mrs_t and mpn_t denote, respectively, the (log) marginal product of labor and the (log) marginal rate of substitution between consumption and leisure.

As illustrated by figure 1, our gap variable can be represented graphically as the vertical distance between the *perfectly competitive* labor supply and labor demand curves, evaluated at the current level of employment (or hours). In much of what follows we assume that our gap variable follows a stationary process with a (possibly nonzero) constant mean, denoted by *gap* (without any time subscript). The latter represents the steady-state deviation between mrs_t and mpn_t. Notice that these assumptions are consistent with both mrs_t and mpn_t being nonstationary, as is likely to be the case in practice as well as in the equilibrium representation of a large class of dynamic business cycle models.

We next relate the gap to the markups in the goods and labor markets. Under the assumption of wage-taking firms, and in the absence of labor adjustment costs, the nominal marginal cost is given by $w_t - mpn_t$, where w_t is (log) compensation per additional unit of labor input (including nonwage costs). Accordingly, we define the aggregate price markup as follows:

$$\mu_t^p = p_t - (w_t - mpn_t) \tag{2}$$

[3] In this respect our results are consistent with recent evidence in Sbordone (2000, 2002), Galí and Gertler (1999), Galí, Gertler and López-Salido (2001), and Christiano, Eichenbaum, and Evans (1997, 1999, 2005) that in somewhat different contexts similarly points to an important role for wage rigidity.

$$= mpn_t - (w_t - p_t). \tag{3}$$

The aggregate wage markup is given by

$$\mu_t^w = (w_t - p_t) - mrs_t, \tag{4}$$

that is, it corresponds to the difference between the wage and the marginal disutility of work, both expressed in terms of consumption. Notice that the wage markup should be understood in a broad sense, including the wedge created by efficiency wages, payroll taxes paid by the firm and labor income taxes paid by the worker, search frictions, and so on.

A variety of frictions (perhaps most prominently, wage and price rigidities) may induce fluctuations in the markups: it is in this respect that these frictions are associated with inefficient cyclical fluctuations, or more precisely, with variations in the aggregate level of (in)efficiency. In particular, given that the marginal rate of substitution is likely to be procyclical, rigidities in the real wage—resulting either from nominal or real rigidities—will give rise to countercyclical movements in the wage markup.[4] Similar rigidities may give rise, in turn, to a countercyclical price markup in response to demand shocks; for holding productivity constant, the marginal product of labor is countercyclical.[5] Alternatively, procyclical movements in competitiveness could induce a countercyclical price markup, as in Rotemberg and Woodford (1996), for example.

To formalize the link between markup behavior and the gap, we first express equation (1) as

$$gap_t = -\{[mpn_t - (w_t - p_t)] + [(w_t - p_t) - mrs_t]\}. \tag{5}$$

Combining equations (3), (4), and (5) then yields a fundamental relation linking the gap to the wage and price markups:

$$gap_t = -(\mu_t^p + \mu_t^w). \tag{6}$$

In the steady state, further,

$$gap = -(\mu^p + \mu^w) < 0, \tag{7}$$

where variables without time subscripts denote steady-state values.

It is natural to assume that $\mu_t^p \geq 0$ and $\mu_t^w \geq 0$ for all t, implying $gap_t \leq 0$ for all t. In this case the level of economic activity is inefficiently low (that is, the gap is always negative), so that (small) increases in our gap measure will be associated with a smaller distortion (that is, an allocation closer to the perfectly competitive one). Notice also that countercyclical movements in these markups imply that the gap is high in booms and low in recessions.

To the extent that we can measure the two markups (or, at least their variation), we can characterize the behavior of the gap, as well as its composition. Constructing our gap variable requires some assumptions about technology and preferences. Below we consider a baseline case with reasonably conventional assumptions. Decomposing the resulting gap variable between wage and price markups requires an additional assumption, namely, that the observed wages used in the construction of the markup reflect the shadow cost of hiring an additional unit of labor. Because this assumption is likely to be more controversial, it is important to keep in mind that it is not necessary in order to measure the gap as a whole, but it is only used in computing its decomposition between the two markups.

Under the assumption of a technology with constant elasticity of output with respect to hours (say, α), we have (up to an additive constant):

$$mpn_t = y_t - n_t, \tag{8}$$

where y_t is output per capita and n_t is hours per capita.[6]

We assume that the (log) marginal rate of substitution for a representative consumer can be written (up to an additive constant) as

$$mrs_t = \sigma c_t + \varphi n_t - \bar{\xi}_t, \tag{9}$$

where c_t is consumption per capita and $\bar{\xi}_t$ is a low-frequency preference shifter. The parameter rameter σ is related to the coefficient of relative risk aversion, and φ measures the curvature of the disutility of labor.[7] Following Hall (1997), we allow for the possibility of low-frequency shifts in preferences over consumption versus leisure, as represented by movements in $\bar{\xi}_t$. These preference shifts may be interpreted broadly to include institutional or demographic changes that affect the labor market, but which are unlikely to be of relevance at business cycle frequencies. We differ from Hall, though, by restricting these shifts to the low frequency. In section IV we provide evidence to justify this assumption.

[4] Models with countercyclical wage markups due to nominal rigidities include Blanchard and Kiyotaki (1987) and Erceg, Henderson, and Levin (2000). Alexopolous (2004) develops a model with a real rigidity due to efficiency wages that can generate a countercyclical wage markup. Alternatively, Hall (1997) stresses the possible role of countercyclical search frictions to account for the behavior of the labor market residual.

[5] With productivity shocks, the markup could be procyclical (because the marginal product of labor moves procyclically in that instance).

[6] Under certain assumptions, that specification is compatible with variable labor utilization—particularly, if labor effort moves roughly proportionately with hours per worker, and the latter is highly positively correlated with aggregate hours (per capita), as the evidence suggests. See, for example, Basu and Kimball (1997) for a detailed discussion.

[7] The parameter φ measures the curvature of the utility function under the standard assumption that labor supply adjusts along the intensive margin (that is, over hours). As we show in appendix A, however, under certain assumptions our framework also allows for labor supply adjustment to occur instead over the extensive margin (that is, over participation). Finally, this log linear representation of the *mrs* has been reconciled with balanced growth in a model with household production (see Baxter and Jermann 1999), or in a generalized indivisible labor model (see King and Rebelo, 1999).

Under the above assumptions our gap variable is thus given by

$$gap_t = (\sigma\ c_t + \varphi\ n_t - \bar{\xi}_t) - (y_t - n_t). \quad (10)$$

Furthermore, we can combine the above assumptions with the definition of the price markup to obtain

$$\mu_t^p = (y_t - n_t) - (w_t - p_t) \quad (11)$$

$$\equiv -s_t. \quad (12)$$

Hence the price markup can be measured (up to an additive constant) as *minus* the (log) *real* unit labor costs, denoted by s_t. Similarly, the wage markup is given by

$$\mu_t^w = (w_t - p_t) - (\sigma\ c_t + \varphi\ n_t) + \bar{\xi}_t. \quad (13)$$

III. The Gap and Its Components: Evidence

We now use the theoretical relations in the previous section to construct measures of the gap and its two main components: the price and wage markups. Our evidence is based on quarterly postwar U.S. data over the sample period 1960:I–2004:IV.[8]

A. *Baseline Case*

Identification of gap and wage markup variations requires that we make an assumption on the coefficient of relative risk aversion σ and on φ, a parameter that corresponds to the inverse of the (Frisch) wage elasticity of labor supply. A vast amount of evidence from microdata suggests a labor supply elasticity mostly concentrated in the range of 0.05–0.5.[9] On the other hand, the business cycle literature tends to use values of unity and higher, using balanced-growth considerations as a justification, as opposed to direct evidence (see, for example Cooley & Prescott, 1995). We use as a baseline value $\varphi = 1$, which we view as a reasonable compromise between the values suggested in the micro and macro literature. In addition, because it will turn out that the costs of fluctuations vary inversely with the Frisch labor supply elasticity, we are biasing our analysis against finding large welfare costs by choosing an elasticity that is above most of the direct estimates in the literature.

The efficiency costs are also increasing in the coefficient of relative risk aversion, because this parameter also affects the steepness of the labor supply curve. There is, however, a similar controversy over the value of this parameter, which equals the inverse intertemporal elasticity of substitution. Direct estimates of the latter tend to fall in the range 0.1–0.3. This evidence suggests a value of σ that varies from 10 to 3.[10] On the other hand, balanced-growth considerations lead the macro literature to a value of unity (again, see Cooley & Prescott, 1995). We will use unity as our baseline case, again opting to bias our parameterization against finding large efficiency costs.

In addition, we need to make an assumption to identify the low-frequency shifter $\bar{\xi}_t$. Let $\tilde{\mu}_t^w \equiv (w_t - p_t) - (\sigma c_t + \varphi\ n_t)$ be the observable component of the wage markup (conditional on values for σ and φ). It follows that

$$\tilde{\mu}_t^w = \mu_t^w - \bar{\xi}_t. \quad (14)$$

From this perspective, the wage markup μ_t^w is the "cyclical" component of $\tilde{\mu}_t^w$, and $\bar{\xi}_t$ is (minus) the "trend" component. We approximate the low-frequency movements of the wage markup by fitting a third-order polynomial of time to $\tilde{\mu}_t^w$.[11]

Finally, before proceeding, we note that the relationships derived in the previous section hold only up to an additive constant. Accordingly, our framework allows us to identify the *variations* over time in the markup and its components, but not their levels. Identification of the level requires that we calibrate the steady-state markup $gap = -(\mu^p + \mu^w)$, an issue to which we turn below.

Our baseline results thus employ measures of the price and wage markups and the gap constructed using, respectively, equations (6), (12), and (13), expressed in terms of deviations from their respective sample means.

Figure 2 presents the time series measure of our gap variable under our baseline assumptions of $\sigma = 1$ and $\varphi = 1$. Notice that this variable comoves strongly with the business cycle, displaying large declines during NBER-dated recessions (represented by the shaded areas in the graph).

We next decompose the movements of the gap into its wage and price markup components. The wage markup measures were constructed using equation (13).[12] The price markup corresponds to minus the log of real unit labor costs, as implied by equation (12). Figure 3 shows the behavior of

[8] The data used to construct the gap variable and its components were drawn from the USECON database commercialized by Estima in Rats format. The time series used (with corresponding mnemonics shown in brackets) include compensation per hour (*LXNFC*), hours all persons (*LXNFH*), and real and nominal output (*LXNFO* and *LXNFI*), all of which refer to the nonfarm business sector. We also make use of the NIPA series for nondurable and services consumption (*CNH* + *GSH*). In addition we also use population over sixteen (*POP16*) to express variables in per capita terms, real GDP (*GDPQ*), implicit GDP deflator (*GDPD*), the Fed funds rate (*FFED*), the spread between the 10-year government bond yield (*FCM10*) and the 3-month Treasury bill rate (*FTB3*), and a commodity price index (*PSCOM*) for our *VAR* exercise in figure 4.

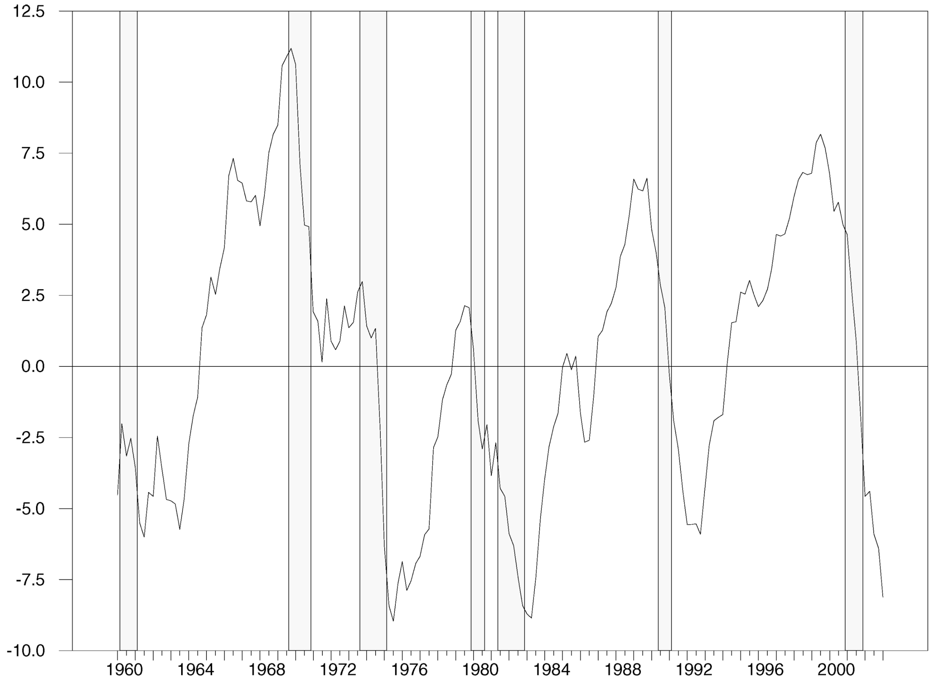
[9] In his survey of the literature, Card (1994) concludes that the intertemporal elasticity of labor supply is "surely no higher than 0.5 and probably no higher than 0.2." However, whether it is appropriate to use the existing micro evidence to calibrate the intertemporal elasticity of labor supply is a matter of considerable controversy, particularly to the extent that employment adjusts along the extensive margin as well as the intensive margin [see, for example, the discussion in Mulligan (1998)].

[10] Using microdata, Barsky et al. (1997) estimate an intertemporal elasticity of substitution of 0.18, implying a coefficient of relative risk aversion slightly above 5. Using macrodata, Hall (1988) concludes that the intertemporal elasticity of substitution ($1/\sigma$) is likely below 0.2.

[11] Because we use the gap measure in subsequent time series analysis, we opt for a high-order polynomial instead of a bandpass filter to detrend the data.

[12] The results are robust to simple adjustments for compositional bias of the real wage, based on Barsky, Parker, and Solon, (1994).

FIGURE 2.—THE GAP: *BASELINE CALIBRATION* ($\sigma = 1$, $\varphi = 1$)



the gap against the wage markup (both relative to their means). To facilitate visual inspection, we plot the inverse wage markup (that is, minus the log wage markup). By definition, the difference between the gap and the inverse wage markup is the inverse price markup. What is striking about the pictures is the strong comovement between the gap and the (inverse) wage markup. Put differently, our evidence suggests that the inefficiency gap seems to be driven largely by countercyclical movements in the wage markup.[13]

[13] As a somewhat cleaner way to illustrate the strong countercyclical relation between the gap and the wage markup, we show later that this pattern also holds conditional on a shock to monetary policy.

FIGURE 3.—THE GAP AND THE WAGE MARKUP: *BASELINE CALIBRATION*

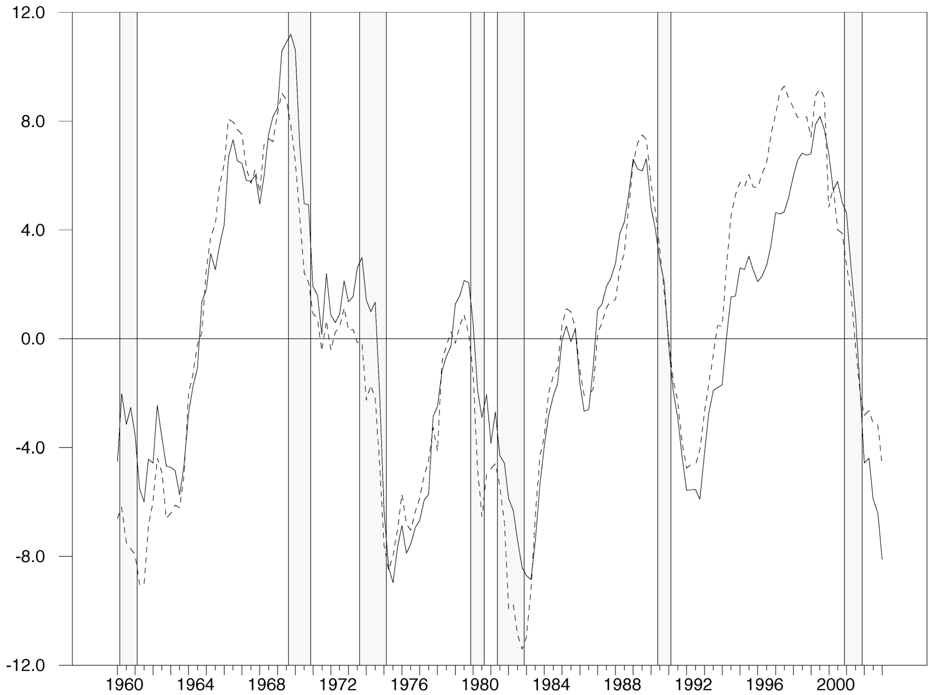


TABLE 1.—BASIC STATISTICS: 1960–2004 — BASELINE CALIBRATION: ($\sigma = 1$, $\varphi = 1$)

			Cross-Correlation			
Variable	SD (%)	ρ	GDP	Gap	Price Mkup	Wage Mkup
GDP	2.6	0.94	1			
Gap	5.1	0.95	0.77	1		
Price markup	2.1	0.88	0.28	−0.02	1	
Wage markup	5.4	0.95	−0.83	−0.92	−0.37	1

Note: Column labeled GDP corresponds to the CBO output gap measure.

To be clear, our conclusion that countercyclical wage markup variation drives the variation in the gap rests on the assumption that wages are allocational and can thus be used to construct a relevant cost measure.[14] Though this assumption is standard in the literature on business cycles and markups (for example, Rotemberg & Woodford, 1999), it is not without controversy. Notice, however, that even if observed wages are not allocational, our gap variable is still appropriately measured, because its construction does not require the use of wage data. Thus our welfare analysis, which depends on the overall gap and not its decomposition, is not affected by this issue.

Table 1 reports some basic statistics that support the visual evidence in figure 3. In particular, the table reports a set of second moments for the gap and its two components (the wage and price markups) and also for detrended (log) GDP, a common indicator of the business cycle. Note first that the percentage standard deviation of the gap is large (relative to detrended output) and that departures of the gap from steady state are highly persistent. In addition, the wage markup is nearly as volatile as the overall gap, and is strongly negatively correlated with the latter, as well as with detrended GDP. This confirms the visual evidence that movements in the gap are strongly associated with countercyclical movements in the wage markup. On the other hand, the price markup is less volatile than the wage markup and does not exhibit a strong contemporaneous correlation with the gap.[15]

B. *Robustness to Alternative Specifications of Technology and Costs*

We next investigate the robustness of our results to the use of alternative specifications of technology and costs. Our baseline case assumes constant elasticity of output with respect to hours and takes the observed average wage as the relevant cost of hiring additional labor. We consider four alternatives to this baseline proposed by Rotemberg and Woodford (1999) in their analysis of cyclical markup behavior. The specific formulations and (subsequent calibrations) we use follow their analysis closely.

Each of the alternatives to the baseline enhances the countercyclical movement in the price markup by making marginal cost more procyclical. The first alternative model assumes a CES production function, thus relaxing the assumption of a constant elasticity of output with respect to labor. The second model allows for overhead labor. The third model, which is based on Bils (1987), allows for the marginal wage to differ from the average wage due to an overtime premium. Finally, the fourth model allows for convex costs of adjusting labor.

In appendix B we present a detailed exposition of how each case affects the measure of the gap and its markup components. We also discuss the calibration. As the appendix makes clear, the CES, overhead labor, and adjustment cost models all alter the marginal product of labor. They accordingly affect the measures of both the overall gap and the price markup. However, they do not affect the measure of the wage markup. On the other hand, the marginal wage model alters only the composition of the gap between the price and wage markups at each point in time, without influencing the gap variable itself (for it affects neither the marginal product of labor nor the household marginal rate of substitution.)

The different panels of table 2 report basic statistics for the alternative measures of the gap and its components, analogous to those reported in table 1 for the baseline case. Overall, the central results from our baseline case are robust to all the alternatives. Both the volatility and persistence of the gap are very similar across all cases. It also remains true in all cases that most of the variation in the gap is due to the variation in the wage markup as opposed to the price markup. The only significant difference is that the price markup tends to display a stronger countercyclical movement than in the baseline case. In contrast to the baseline model, the price markup in the CES, overhead labor, and marginal wage models is negatively correlated with the output gap. In all the alternative models, further, the negative comovement of the price markup with our gap variable is larger than in the baseline case.

In panel A of figure 4 we show that the historical movement in the gap is robust to the alternative cases. We plot the time series of the gap for the baseline case against all the alternatives except the marginal wage model (for which the

[14] Some indirect evidence that wages are allocational is found by Sbordone (2002) and Galí and Gertler (1999), who show that firms appear to adjust prices in response to measures of marginal cost based on wage data. In turn, as shown in Galí (2001), they do not respond to marginal cost measures that employ the household's marginal rate of substitution in place of the wage, as would be appropriate if wages were not allocational.

[15] However, the relatively weak comovement of the price markup with detrended output is useful for understanding the dynamics of inflation and the recent evidence on the New Keynesian Phillips curve. See Sbordone (2002) and Galí and Gertler (1999).

TABLE 2.—BASIC STATISTICS: 1960–2004 — ROBUSTNESS: ALTERNATIVE MEASURES OF REAL MARGINAL COST

			Cross-Correlation		
Variable	SD (%)	ρ	GDP	Gap	Price Mkup
Gap:					
CES	4.8	0.95	0.72		
Overhead	6.4	0.95	0.80		
Adj. cost	5.3	0.92	0.81		
Bils	5.1	0.95	0.77		
Price markup:					
CES	2.0	0.92	−0.02	−0.21	
Overhead	2.5	0.90	−0.21	−0.54	
Adj. cost	2.1	0.78	0.13	−0.24	
Bils	3.0	0.92	−0.17	−0.45	
Wage markup:					
CES	4.8	0.94	−0.71	−0.92	−0.20
Overhead	5.5	0.94	−0.83	−0.93	0.20
Adj. cost	5.5	0.94	−0.83	0.93	−0.25
Bils	4.4	0.94	−0.74	−0.84	−0.05

Note: See table 1.

measure of the overall gap is the same as for the baseline). Clearly, the gap measures move very tightly together in all cases. Finally, though the gap measure in the marginal model is the same as in the baseline case, the division into price and wage markup movements differs. Accordingly, in panel B of figure 4 we plot the wage markup for the marginal model (Bils adjustment) relative to the baseline case. As the figure shows, the broad pattern in the movement of the wage markup is very similar across the two cases.

To summarize: the results thus far suggest that the business cycle is associated with large coincident movements in the efficiency gap. Thus, under our framework, the evidence suggests that countercyclical markup behavior is potentially an important feature of the business cycle. A decomposition of the gap further suggests that the countercyclical movement in the wage markup is by far the most important source of overall variations in the gap. Thus, to the extent that wages are allocational, some form of wage rigidity, either real or nominal, may be central to business fluctuations.

IV. Labor Supply Shifts and the Gap

We have proceeded under the interpretation that our measured gap between the marginal rate of substitution and the marginal product of labor reflects countercyclical markup behavior. In his baseline identification scheme, however, Hall modeled this gap as an unobserved preference shock, though he was clear to state that he did not take this hypothesis literally. Subsequent literature, however (for example, Holland & Scott, 1998; Francis & Ramey, 2001; Uhlig, 2002) has indeed interpreted this residual as reflecting either exogenous labor supply shifts or some other unspecified exogenous driving force. In this section we show that the high-frequency movements in the gap cannot be simply due to exogenous preference shifts. Rather, the evidence is instead compatible with our countercyclical markup interpretation.

Let us follow Hall (1997) in assuming that the marginal rate of substitution is now augmented with a preference shock ξ_t that contains a cyclical component $\tilde{\xi}_t$, as well as a trend component $\bar{\xi}_t$:

$$mrs_t = c_t + \varphi\, n_t - \xi_t \tag{15}$$

with

$$\xi_t = \bar{\xi}_t + \tilde{\xi}_t,$$

where we maintain our baseline assumption that the coefficient of relative risk aversion, σ, is unity. Hall then defines the residual x_t as the difference between the "observable" component of the marginal rate of substitution, $c_t + \varphi\, n_t$, and the marginal product of labor, $y_t - n_t$:

$$x_t \equiv (c_t + \varphi\, n_t) - (y_t - n_t). \tag{16}$$

The issue then is how exactly to interpret the movement in Hall's residual. Using the augmented specification (15) of the marginal rate of substitution allowing for preference shocks, together with equation (8) and the definition (1) of the inefficiency gap, it is possible to express x_t as follows:

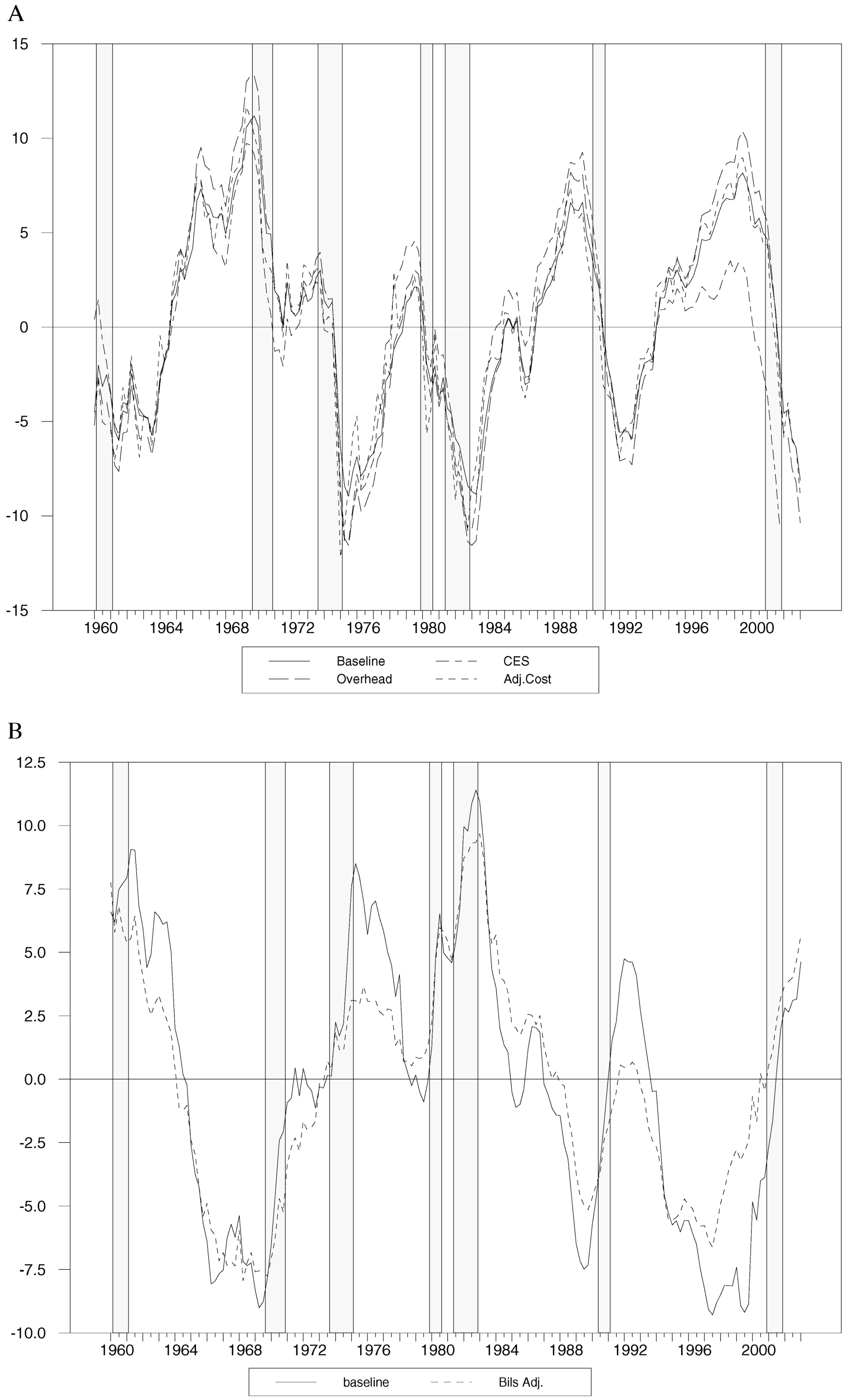
$$x_t \equiv (mrs_t - mpn_t) + \xi_t \tag{17}$$

$$= -(\mu_t^p + \mu_t^w) + \xi_t. \tag{18}$$

Hall's assumption of perfect competition in both goods and labor markets implies $\mu_t^p = \mu_t^w = 0$. This allows him to interpret the variable x_t as a preference shock, for under this assumption $x_t = \xi_t$.[16] Notice that under these circumstances the efficiency gap is 0, as there are no imperfections in either goods or labor markets. On the other hand, if preferences are not subject to shocks ($\xi_t = 0$, all t), and we allow

[16] See also Baxter and King (1991). Holland and Scott (1998) construct similar measures for the United Kingdom.

FIGURE 4.—(A) THE GAP: ALTERNATIVE MEASURES, *BASELINE CALIBRATION* ($\sigma = 1$, $\varphi = 1$), (B) WAGE MARKUP: BILS ADJUSTMENT



for departures from perfect competition, x_t will purely reflect movements in markups, that is, $x_t = -(\mu_t^p + \mu_t^w)$. In the latter instance, x_t corresponds exactly to our inefficiency gap, that is, $x_t = gap_t$, for all t.

Note that if x_t indeed reflects exogenous preference shocks, it should be invariant to any other type of disturbance. In other words, the null hypothesis of preference shocks implies that x_t should be exogenous. We next present

TABLE 3.—GRANGER-CAUSALITY TESTS (1960–2004): BASELINE CALIBRATION ($\sigma = 1$, $\varphi = 1$), BIVARIATE VAR (4 LAGS)

Variable	*p*-Value				
	Baseline	CES	Overhead	Adj. cost	Bils
CBO output gap	0.000	0.000	0.000	0.004	0.000
Nominal interest rate	0.270	0.048	0.200	0.106	0.275
Yield spread	0.006	0.000	0.003	0.003	0.006

Note: The values reported are *p*-values for the null hypothesis of no Granger causality from each variable listed to Hall x (F-test). Filtered data using third-order polynomial in time.

two tests that reject the null of exogeneity, thus rejecting the preference shock hypothesis.

First, we test the hypothesis of no Granger causality from a number of variables to our gap measure. The variables used are: detrended GDP, the nominal interest rate, and the yield spread. Both the nominal interest rate and the yield spread may be thought of as rough measures of the stance of monetary policy, whereas detrended GDP is just a simple cyclical indicator. Table 3 displays the *p*-values for several Granger-causality tests. These statistics correspond to bivariate tests using alternative lag lengths. They indicate that the null of no Granger causality is rejected for all specifications, at conventional significance levels. That finding is robust to reasonable alternative calibrations of σ and φ. Overall, the evidence of Granger causality is inconsistent with the hypothesis that x_t mainly reflects variations in preferences.

As a second test, we estimate the dynamic response of our gap variable to an identified exogenous monetary policy shock. The identification scheme is similar to the one proposed by Christiano, Eichenbaum, and Evans (1999) and others. It is based on a VAR that includes measures of output, the price level, commodity prices, and the Federal Reserve Funds rate, to which we add our gap measure (or, equivalently, Hall's residual) and the price markup. From the gap and the price markup response we can back out the behavior of the wage markup, using equation (6). We identify the monetary policy shock as the orthogonalized innovation to the Federal Reserve Funds rate, under the assumption that this shock does not have a contemporaneous effect on the other variables in the system.

Figure 5 shows the estimated responses to a contractionary monetary policy shock. The responses of the nominal rate, output, consumption, and prices are similar to those found in Christiano et al. (1999), Bernanke and Mihov (1998), and other papers in the literature. Most interestingly for our purposes, the inefficiency gap declines significantly in response to the unanticipated monetary tightening. Its overall pattern of response closely mimics the response of output. This endogenous reaction, of course, is inconsistent with the preference shock hypothesis, but fully consistent with our hypothesis that countercyclical markups may underlie the cyclical variation in the Hall residual. In this connection, note that the monetary shock induces a rise in the wage markup that closely mirrors the decline in the gap, both in the shape and in the magnitude of the response. This countercyclical movement in the wage markup is consistent with evidence on unconditional comovements presented in table 1. The price markup also rises, though with a significant lag. Apparently, the sluggish response of wages, which gives rise to a strong countercyclical movement in the wage markup, delays the rise in the price markup.[17] In any event, the decline in the inefficiency gap is clearly associated with a countercyclical rise in markups.

To be clear, because preference shocks are not observable, it is not possible to directly determine the overall importance of these disturbances. Although our evidence rejects the hypothesis that exogenous preference variation drives all the movement in our gap measure, it cannot rule out the possibility that some of this movement is due to preference shocks. Yet, to the extent that preference shocks are mainly a low-frequency phenomenon, they are likely to be captured by the trend component associated with our low-frequency filter (together with other institutional and demographic factors which may lead to low-frequency variations in markups). In this instance our filtered gap series, which isolates the high-frequency movement in this variable, is likely to be largely uncontaminated by exogenous preference variations.

V. Welfare and the Gap

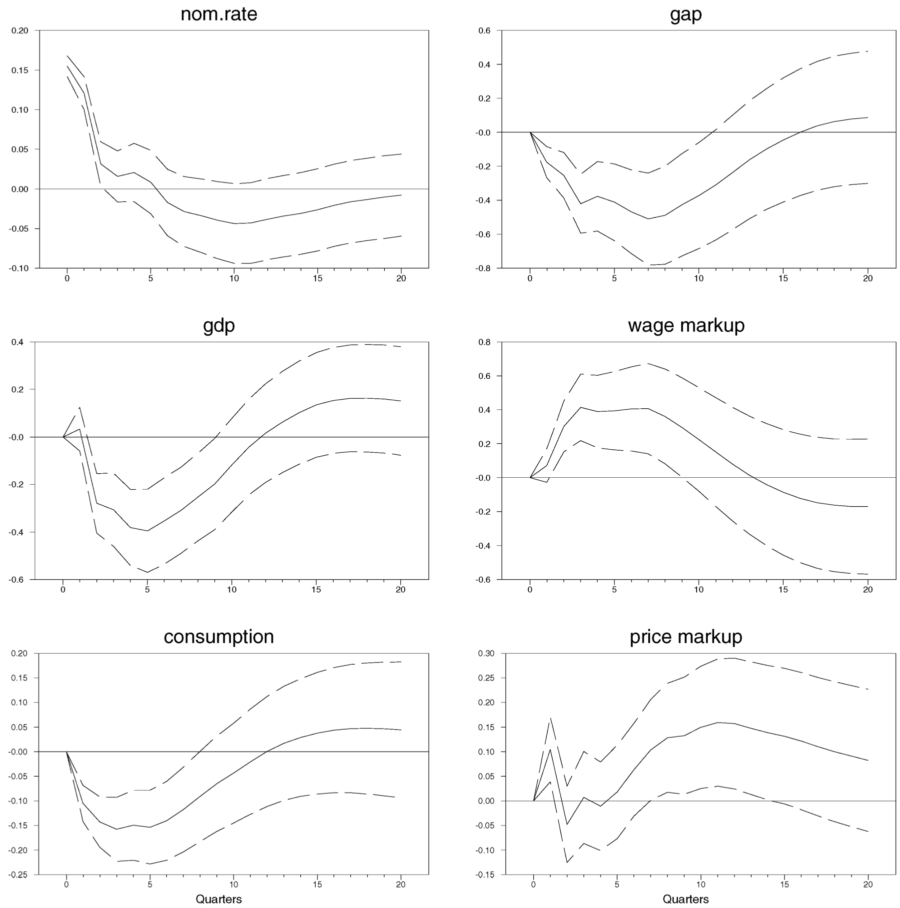
We next propose a simple way to measure the welfare costs of fluctuations in the degree of inefficiency of aggregate resource allocations, as captured by our gap variable. We then apply this methodology to postwar U.S. data. In addition to obtaining a measure of the average cost of gap fluctuations, we also compute the welfare losses during particular episodes, including the major postwar recessions.

As we noted in the introduction, our approach differs from Lucas (1987) and others in focusing on the costs stemming from fluctuations in the degree of inefficiency of the aggregate resource allocation, as reflected by the movements in our gap variable.[18] As in Ball and Romer (1987), the cycle generates losses on average within our framework

[17] As Galí and Gertler (1999) and Sbordone (2002) observe, the sluggish behavior of the price markup helps explain the inertial behavior of inflation, manifested in this case by the delayed and weak response of inflation to the monetary shock. Staggered pricing models relate inflation to an expected discounted stream of real marginal costs, which corresponds to the reciprocal of the price markup. The sluggish response to the price markup translates into sluggish behavior of real marginal cost.

[18] For other approaches to measuring the unconditional costs of fluctuations see, for example, Barlevy (2004) and Beaudry and Pages (2001). For a very early attempt to measure the welfare cost of inefficiently high unemployment, see Gordon (1973).

FIGURE 5.—DYNAMIC EFFECTS OF MONETARY POLICY SHOCKS: *BASELINE CALIBRATION, SAMPLE PERIOD 1960–2004*



Note: 95% confidence bands for impulse responses are based on 5000 Monte Carlo replications. Sample period: 1960:1–2004:3.

because the welfare effects of employment fluctuations about the steady state are asymmetric. As figure 1 illustrates, given that the steady-state level of employment is inefficient (due to positive price and wage markups in the steady state), the efficiency costs of an employment contraction below the steady state will exceed the benefits of a symmetric increase. In particular, note that the vertical distance between the labor demand and supply curves rises as employment falls below the steady state, and falls as it rises above. The quantitative effect of this nonlinearity on the welfare cost of fluctuations ultimately depends on the slopes of the labor demand and supply curves, and on the steady-state distance relative to the first-best, perfectly competitive steady state.

Underlying this measure of the average costs of fluctuations are the gross gains from booms and losses from recessions. As we elaborate, under our maintained hypothesis that the flexible price equilibrium is distorted (due to imperfect competition and taxes, and so on), there are significant first-order welfare losses from employment contractions below the steady state, as well as gains from movements above. Below we present a time series measure of these gross efficiency costs and benefits, along with an overall net measure.

A. *A Welfare Measure*

We now proceed to derive our welfare measure. The economy is assumed to fluctuate around an underlying "frictionless" path characterized by a constant gap level:

$$GAP = \frac{\overline{MRS_t}}{\overline{MPN_t}} = \exp\{-\mu\} \equiv 1 - \Phi < 1,$$

where upper bars denote values along a constant gap path, and μ is (minus) the steady-state value of our (log) gap variable. A second-order approximation of the period utility around its level along the underlying constant-gap path yields

$$\Delta_t \equiv U(C_t,N_t) - U(\bar{C}_t,\bar{N}_t) = \bar{U}_{c,t}\bar{C}_t\left(\tilde{c}_t + \frac{1-\sigma}{2}\tilde{c}_t^2\right) + \bar{U}_{n,t}\bar{N}_t\left(\tilde{n}_t + \frac{1+\varphi}{2}\tilde{n}_t^2\right),$$

where the tildes denote log deviations from the underlying constant-gap path, that is, $\tilde{x}_t \equiv \log(X_t/\bar{X}_t)$, and where $\varphi \equiv -(\bar{U}_{nn,t}\bar{N}_t)/\bar{U}_{n,t}$ and $\sigma \equiv -\bar{U}_{cc,t}\bar{C}_t/\bar{U}_{c,t}$.

In order to maintain tractability, we make two additional assumptions. First, we assume that all output is consumed, which in turn implies $\tilde{c}_t = \tilde{y}_t$ for all t. Secondly, we assume that output is linearly related to hours in equilibrium, that is, $y_t = a_t + n_t$, thus implying $\tilde{n}_t = \tilde{y}_t$. The latter assumption is consistent with the notion that variations in the stock of capital are negligible at business cycle frequencies, and that the rate of capital utilization is proportional to hours. Notice that these two assumptions imply that

$$-\frac{\bar{U}_{n,t}\bar{N}_t}{\bar{U}_{c,t}\bar{C}_t} = 1 - \Phi.$$

Hence, we can rewrite the second-order approximation as

$$\Delta_t = \bar{U}_{c,t}\bar{C}_t\left(\Phi\tilde{y}_t - \frac{1}{2}[(\sigma + \varphi) - (1 - \Phi)(1 + \varphi)]\tilde{y}_t^2\right). \tag{19}$$

Furthermore, under the previous assumptions, together with the log linear specification for the marginal rate of substitution in equation (9), it is easy to check that

$$\widehat{gap}_t = (\sigma + \varphi)\tilde{y}_t,$$

where $\widehat{gap}_t \equiv gap_t - gap$. Using the previous expression to substitute for $\tilde{y}_t$ in equation (19), we obtain

$$\frac{\Delta_t}{\bar{U}_{c,t}\bar{C}_t} = \frac{1}{\sigma + \varphi}(\Phi\widehat{gap}_t - \psi\widehat{gap}_t^2) \equiv \omega(\widehat{gap}_t) \tag{20}$$

where $\psi \equiv \frac{1}{2}\left[1 - \frac{(1 - \Phi)(1 + \varphi)}{\sigma + \varphi}\right]$.

Notice that $\omega(\widehat{gap}_t)$ is the period efficiency loss or gain from the gap's deviations from its steady-state value, expressed as a percentage of the frictionless level of consumption $\bar{C}_t$. The first term in the parentheses, the linear term, reflects the symmetric first-order costs and benefits from the gap moving below and above the steady state, due to the positive steady-state markup μ (implying $\phi > 0$). The quadratic term captures the asymmetric, second-order effects of gap fluctuations on welfare. For plausible values of μ, σ, and φ we have $\psi > 0$. In that case ω is concave, implying that a reduction in the gap below its steady-state value results in an efficiency loss that exceeds the gain stemming from a commensurate increase in the gap above its steady state.

We can use equation (20) to calculate a time series of the efficiency gain or loss in each quarter t. To obtain a measure of the average welfare cost over time analogous to those found in the literature, we take the unconditional expectation of equation (20) to obtain

$$E\left\{\frac{\Delta_t}{\bar{U}_{c,t}\bar{C}_t}\right\} = -\frac{\psi}{\sigma + \varphi}\,\mathrm{var}(gap_t), \tag{21}$$

where $\mathrm{var}(gap_t)$ is the variance of our gap measure. Notice that, as a result of the concavity of ω, the expected welfare effects of fluctuations in the gap variable are negative, that is, these fluctuations imply *losses* in expected welfare. This loss, further, is of second order, as it is linearly related to the variance of the inefficiency gap. It is, however, potentially large, depending in particular on the magnitude $\mathrm{var}(gap_t)$. As section 3 suggests, $\mathrm{var}(gap_t)$ is potentially large if the labor supply is inelastic or the risk aversion is high.

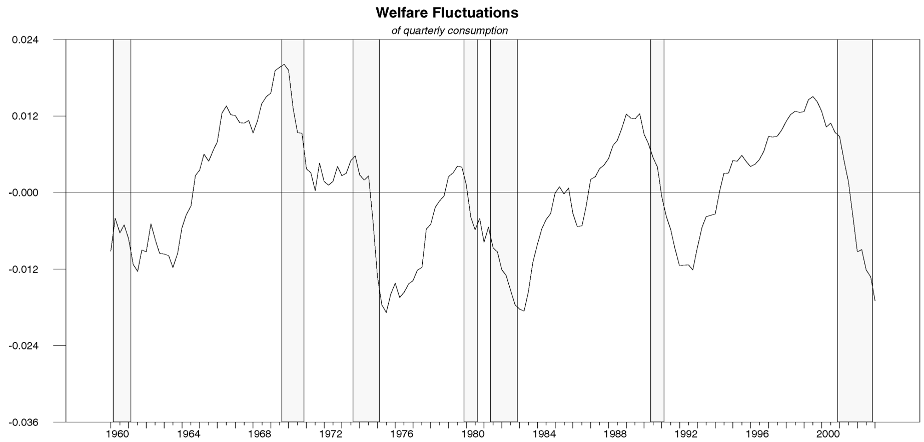
To be clear, our approach provides a lower bound on the measure of the total welfare costs of fluctuations. The reason is simple: it does not include the welfare costs from *efficient* fluctuations in consumption and employment. Suppose, for example, that the data were generated by a real-business-cycle model with frictionless, perfectly competitive markets. We should then expect to see no variation in our gap measure, as the resource allocation would always be efficient. Our metric would then indicate no welfare costs of fluctuations, but some losses would still be implied by the variability of consumption and leisure (under standard convexity assumptions on preferences). It is also important to stress that, to the extent that the steady-state value of the gap corresponds also to the average value around which the economy fluctuates (as assumed above), *average* welfare losses will only be of second order. On the other hand, our efficiency cost measure suggests possible first-order effects at any moment in time: As reflected in equation (20) and illustrated further below, deviations in the gap variable from that steady state may have nonnegligible first-order welfare effects, with the gap declines associated with recessions generating large welfare losses.

B. *Some Numbers*

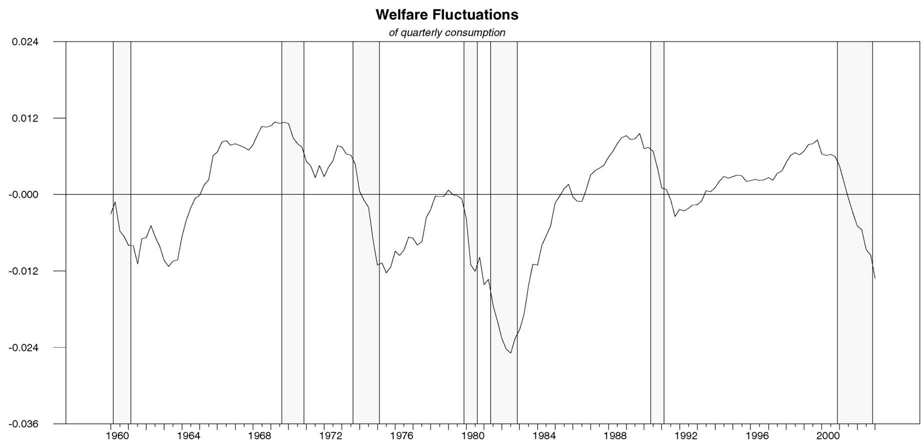
Equation (20) provides a real-time measure of the efficiency costs of deviations of our gap variable from steady state. Accordingly, we construct a quarterly time series of $\omega(\widehat{gap}_t)$, taking as input our measure of the gap. We consider three different parameterizations: first, our baseline case with $\sigma = 1$ and $\varphi = 1$; second, a case where we raise the risk aversion with $\sigma = 5$ and $\varphi = 1$; and third, a case where we reduce the labor supply elasticity with $\sigma = 1$ and $\varphi = 5$ (implying a Frisch labor supply elasticity of 0.2). For the parameter μ, the sum of the steady-state wage and price markups, we assume a value of 0.50. A value of 0.15 to 0.20 is plausible for the steady-state price markup (see Rotemberg and Woodford, 1999). Inasmuch as the steady-state wage markup depends on tax distortions as well as workers' market power, 0.30 to 0.35 seems a reasonable lower bound given the evidence on average labor tax rates. This range is also roughly consistent with the evidence in Mulligan (2002).

FIGURE 6.—THE WELFARE EFFECTS OF POSTWAR U.S. FLUCTUATIONS

(σ=1, φ=1, μ=0.50)



(σ=5, φ=1, μ=0.50)



(σ=1, φ=5, μ=0.50)

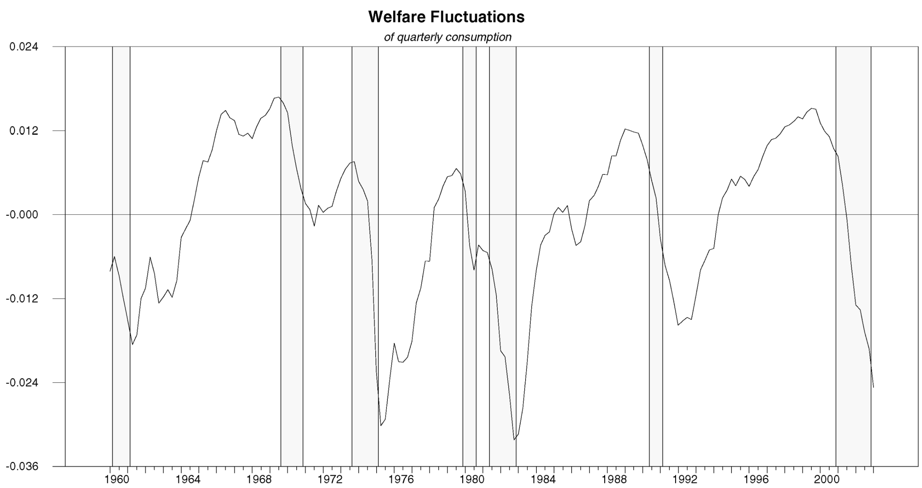


Figure 6 plots the resulting time series over the sample 1960:4–2004:4. The value at each period t can be interpreted as the efficiency gain or loss in percentage units of consumption associated with the deviation of the inefficiency gap from its steady state. Our baseline parametrization indicates substantial fluctuations in welfare resulting

Table 4.—Welfare Costs of Fluctuations (1960–2004)

	Percentage of One Year's Consumption	
σ	$\varphi = 1$	$\varphi = 5$
1	0.010	0.043
5	0.027	0.059
10	0.049	0.080

Note: Based on calibration $\mu = 0.5$. The data were filtered using a third-order polynomial in the time. Welfare computations cover the sample period 1960:1–2004:3.

Table 5.—The Welfare Costs of Recession Episodes

		Percentage of One Year's Consumption		
σ	φ	1970s	1980s	1990s
1	1	−4.58	−4.69	−2.26
1	5	−6.18	−6.37	−3.22
5	1	−2.88	−7.23	−0.39
5	5	−4.89	−8.00	−1.65

Note: See table 4.

with changes in the degree of aggregate efficiency. For example, efficiency-based welfare losses during the major recessions are on average around 2.0% of period consumption around the time of the respective troughs. Furthermore, during the major recessions these large welfare losses tend to persist for a number of years. Likewise, the average gain at the major cyclical peaks is a bit over 1.0%. These gains also tend to persist.

With higher risk aversion ($\sigma = 5$) or lower labor supply elasticity ($\varphi = 5$) the losses during downturns go up while the gains during booms decline. In either case the labor supply curve is steeper than in the baseline case, enhancing the asymmetric effects on efficiency of symmetric movements in employment above and below its natural level. In the case of low labor supply elasticity, for example, the efficiency losses during the major recessions hover around 3.0% of consumption per period around the time of the respective troughs.

In table 4 we present a measure of the average welfare cost of the cycle, based on equation (21). As we noted earlier, the measure is simply proportional to the square of the gap. We construct estimates for alternative values of the parameters φ and σ. For the parameterization that corresponds to our baseline case ($\varphi = 1$, $\sigma = 1$), we estimate the average efficiency costs of postwar U.S. business fluctuations to be quite small, roughly 0.01% of steady-state consumption. In this case, the asymmetric movements in efficiency over the cycle are small, implying that the gains during booms seen in figure 6 approximately cancel the losses during recessions. The estimates of efficiency losses go up as we reduce labor supply elasticity and increase risk aversion. With $\varphi = 5$ and $\sigma = 10$, for example, the average efficiency costs go up to 0.08. This number, however, is still not large and is in the range of Lucas's original estimates.

Any measure of the average cost of business cycles, however, obscures the fact that individual recessionary episodes may be rather costly. What moderates the impact of these episodes on the overall welfare measure is the fact they have been infrequent, particularly over the last several decades. One reason for this may be that stabilization policy has been reasonably effective. Another possibility is that the economy has been subject to smaller shocks. In either event, it is of interest to examine efficiency losses during the major recessionary episodes. Doing so provides a sense of the gains from avoiding future recessions (either by good policy or by good luck).

There are three distinct recessionary periods in our sample where the gap fell below the steady state prior to the trough and then returned to steady state following the trough. These periods include the two major recessions of the mid-1970s and of the early 1980s and also the recession of the early 1990s.[19] For each recessionary period, we report the cumulative efficiency losses over the recession as a percentage of one year's consumption. We again consider a variety of parameterizations, including our baseline case.

Table 5 reports the efficiency losses for the three recessionary periods. For our baseline case ($\sigma = 1$ and $\varphi = 1$), the efficiency costs of each of the two major recessions was large, roughly 4.5% of one year's consumption. For the milder recession of the 1990s the cost was still nontrivial, more than 2% of one year's consumption. With lower labor supply elasticity ($\varphi = 5$), the efficiency costs of the two major recessions rise to over 6.0% of one year's consumption, and the the cost of the 1990s recession rises to over 3.0%.

Increasing risk aversion boosts the costs of the recession in the early 1980s. With $\sigma = 5$, the efficiency cost of the downturn rises to over 7.0% of steady state consumption in the case of high labor supply elasticity ($\varphi = 1$). It goes up to 8.0% when combined with low labor supply elasticity ($\varphi = 5$). Interestingly, for the other two recessions, raising risk aversion actually tends to reduce the estimated efficiency cost. Intuitively, higher risk aversion places more weight on consumption in the measure of gap fluctuations. Because the decline in consumption was relatively modest in each of these downturns, as compared to the 1980–1982 recession, increasing risk aversion tends to dampen gap fluctuations over these periods. For these reasons it reduces measured efficiency losses.

Overall, our results suggest only modest *average* efficiency losses from fluctuations. However, major recessionary episodes appear to entail rather significant losses.

VI. Concluding Comments

At the risk of considerable oversimplification, it is possible to classify modern business cycle models into two types. The first class attempts to explain quantity fluctuations by appealing to high degrees of intertemporal substi-

[19] For the other recessions in the sample (the early 1960s and the early 2000s), we do not have the complete swing of the gap below and back to steady state.

tution in an environment of frictionless markets. The second instead appeals to countercyclical markups owing to particular market frictions. In this regard, there has been a considerable debate as to whether the markup is indeed countercyclical [see Rotemberg and Woodford (1999) for a summary]. Much of this debate has been centered around price markup measures that use wage data to calculate the cost of labor. We show, however, that the markup is highly countercyclical, using the household's consumption leisure tradeoff as the shadow cost of labor, as theory would suggest. Under this identification scheme, the markup corresponds exactly to the labor market residual studied by Hall (1997) and others. Whether the countercyclical markup variation is driven primarily by product market or labor market behavior is, however, an open question. To the extent that wages are allocative, we find that labor market frictions are the key factor. As we discussed, however, the exact form that these frictions may take (such as nominal wage rigidity, efficiency wages, or search frictions) is also an open question.

A second message of this paper is that to the extent that our markup interpretation of the efficiency gap is correct, business cycles may involve significant efficiency costs. To be sure, our results suggest that these efficiency losses are modest when averaged over time. This result occurs, however, because—whether by good luck or good policy—significant recessions have not often occurred since World War II. We find, however that when they do occur, the efficiency costs may indeed be quite large. These results obtain for reasonably standard assumptions on preferences (for example, a coefficient of relative risk aversion of 5 and a unit-elasticity Frisch labor supply). Thus, though the gains from eliminating all fluctuations may not be large—as suggested by the existing literature—there nonetheless do appear to be significant efficiency benefits from avoiding severe recessions.

Finally, we observe that our calculation ignores at least several important considerations that might be leading us to understate the efficiency costs of recessions. First, within our framework, a reduction in hours leads to increased enjoyment of leisure, which partially offsets the impact of the output decline. In reality, workers who are laid off during recessions do not simply get to enjoy the time off, but rather have to look for a new job. In addition, there is often a loss of human capital that was specific to the previous employer. Second, our calculation ignores the costs of fluctuations in price and wage inflation associated with variations in markups resulting from nominal rigidities (see, for example, Woodford, 1999). For this reason, our metric may overstate the gains from booms (and understate the losses from recessions). To the extent that the costs of high inflation roughly offset the efficiency gains from the boom, our measure of the gross efficiency loss of the recession may provide a more accurate indicator of the costs of these episodes. Taking these considerations into account is on the agenda for future research.

REFERENCES

Alexopolous, Michelle, "Unemployment and the Business Cycle," *Journal of Monetary Economics* 51 (2004), 277–298.

Ball, Lawrence, and David Romer, "Are Prices Too Sticky?" *Quarterly Journal of Economics* 104 (1987), 507–524.

Barlevy, Gadi, "The Cost of Business Cycles under Endogenous Growth," *American Economic Review* 94:4 (2004), 964–990.

Barsky, Robert, Jonathon Parker, and Gary Solon, "Measuring the Cyclicality of Real Wages: How Important is Composition Bias?" *Quarterly Journal of Economics* 109 (1994), 1–25.

Barsky, Robert, F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro, "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *Quarterly Journal of Economics* 112 (1997), 537–579.

Basu, S., and M. Kimball, "Cyclical Productivity with Unobserved Input Variation," NBER working paper no. 5915 (1997).

Baxter, M., and Robert King, "Productive Externalities and Business Cycles," Federal Reserve Bank of Minneapolis, Institute for Empirical Macroeconomics, discussion paper no. 53 (1991).

Baxter, M., and U. J. Jermann, "Household Production and the Excess of Sensitivity of Consumption to Current Income," *American Economic Review* 89:4 (1999), 902–920.

Beaudry, Paul, and Carmen Pages, "The Cost of Business Cycles and the Stabilization Value of Unemployment Insurance," *European Economic Review* 45:8 (2001), 1545–1572.

Bernanke, Ben, and Ilyan Mihov, "Measuring Monetary Policy," *Quarterly Journal of Economics* 113 (1998), 869–902.

Blanchard Olivier, J., and Nobuhiro Kiyotaki, "Monopolistic Competition and the Effects of Aggregate Demand," *American Economic Review* 77 (1987), 647–666.

Card, David, "Intertemporal Labor Supply: An Assessment" (pp. 49–78), in Christopher Sims (Ed.), *Advances in Econometrics: Sixth World Congress,* Vol. II, (Cambridge University Press, 1994).

Chari, V. V., Kehoe, P., and Ellen R. McGrattan, "Business Cycles Accounting," Federal Reserve Bank of Minneapolis staff report 328 (2004).

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans, "Sticky Prices and Limited Participation Models: A Comparison," *European Economic Review* 41 (1997), 1201–1249.

——— "Monetary Policy Shocks: What Have We Learned and to What End?" in J. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics,* Vol. 1A (Amsterdam: North-Holland, 1999).

——— "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy* 113:1 (2005), 1–45.

Cooley, Thomas, and Edward Prescott, "Economic Growth and Business Cycles," in Thomas Cooley (Ed.), *Frontiers of Business Cycle Research* (Princeton University Press, 1995).

Eichenbaum, Martin, Lars Hansen, and Kenneth Singleton, "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice under Uncertainty," *Quarterly Journal of Economics* 103:1 (1988), 51–78.

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

Francis, Neville, and Valerie Ramey, "Is the Technology-Driven Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited," UCSD mimeograph (2001).

Galí, Jordi, "The Case for Price Stability. A Comment," in A. Garcia-Herrero, V. Gaspar, L. Hoogduin, J. Morgan, and B. Winkler (Eds.), *Why Price Stability?* (Frankfurt am Main, European Central Bank, 2001).

Galí, J., and Mark Gertler, "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics* 44 (1999), 195–222.

Galí, J., Mark Gertler, and J. David López-Salido, "European Inflation Dynamics," *European Economic Review* 45:7 (2001), 1121–1150.

——— "Markups, Gaps, and the Welfare Costs of Business Fluctuations," NBER working paper no. 8850 (2002).

Gordon, Robert J., "The Welfare Cost of Higher Unemployment," *Brookings Papers on Economic Activity* 4:1 (1973), 133–195.

Hall, Robert E., "Intertemporal Substitution in Consumption," *Journal of Political Economy* 96:2 (1988), 339–357.
——— "Macroeconomic Fluctuations and the Allocation of Time," *Journal of Labor Economics* 15:1, pt. 2 (1997), S223–S250.
Holland, A., and A. Scott, "Determinants of U.K. Business Cycles," *Economic Journal* 108 (1998), 1067–1092.
King, Robert, and Sergio Rebelo, "Resuscitating Real Business Cycles," in J. Taylor and M. Woodford (Eds.), *The Handbook of Macroeconomics,* Vol. 1B (Amsterdam North-Holland, 1999).
Lucas, Robert E. *Models of Business Cycles* (Oxford University Press, 1987).
——— "Macroeconomic Priorities," *American Economic Review* 93:1 (2003), 1–14.
Mulligan, Casey B., "Substitution over Time: Another Look at Life-Cycle Labor Supply," in Ben Bernanke and Julio Rotemberg (Eds.), *NBER Macroeconomics Annual,* Vol. 13 (Cambridge, MA: MIT Press, 1998).
——— "A Century of Labor-Leisure Distortions," NBER working paper no. 8774 (2002).
Pencavel, John, "The Market Work Behavior and Wages of Women, 1975–94," *Journal of Human Resources* 33 (1998), 771–804.
Rotemberg, Julio, and Michael Woodford, "The Cyclical Behavior of Prices and Costs," in J. Taylor and M. Woodford (Eds.), *The Handbook of Macroeconomics,* Vol. 1B (Amsterdam: North-Holland, 1999).
Sbordone, Argia M. "An Optimizing Model of Wage and Price Dynamics," Rutgers University mimeograph (2000).
——— "Prices and Unit Labor Costs: Testing Models of Pricing Behavior," *Journal of Monetary Economics* 45:2 (2002), 265–292.
Uhlig, Harald, "What Moves Real GNP?" presented at CEPR-ESSIM Conference, May Tarragona, Spain (2002).
Woodford, M., "Inflation Stabilization and Welfare," NBER working paper no. 5684 (1999).

APPENDIX A

The Household's MRS

Here we illustrate that the expression we use for the household's marginal rate of substitution between consumption and leisure, equation (9), may be motivated either by making the standard assumption that labor supply adjusts along the intensive margin or, under certain assumptions, that the adjustment is along the extensive margin. Our argument is based on Mulligan (1998).

Case I. Labor Supply Adjustment along the Intensive Margin: Let C_t and N_t denote consumption and hours worked, respectively. Assume a representative agent with preferences given by

$$\frac{1}{1-\sigma} C_t^{1-\sigma} - \frac{1}{1+\varphi} N_t^{1+\varphi}.$$

It follows that

$$MRS_t = C_t^{\sigma} N_t^{\varphi}.$$

By taking the log of each side of this relation we obtain equation (9).

Case II. Labor Supply Adjustment along the Extensive Margin: Now assume that individuals either do not work or work a fixed amount of hours per week. Suppose there is a representative household with a continuum of members represented by the unit interval, and who differ according to their disutility of work. Specifically, let us assume that j^{φ} is the disutility of work for member j. Under perfect consumption insurance within the household, and interpreting N_t as the fraction of working household members in period t, total household utility will be given by

$$\frac{1}{1-\sigma} C_t^{1-\sigma} - \int_0^{N_t} j^{\varphi}\, dj.$$

Note that

$$\int_0^{N_t} j^{\varphi}\, dj = \frac{1}{1+\varphi} N_t^{1+\varphi}$$

Accordingly, the utility function for the family in this case is isomorphic to the case of adjustment along the intensive margin. It follows that the marginal rate of substitution has the same form as well.

APPENDIX B

Alternative Specifications

We now present the details that underlie the alternative measures of the gap and its components that we examined in the text. Our baseline case assumes constant elasticity of output with respect to hours and takes the observed average wage as the relevant cost of hiring additional labor. The four alternatives we consider are those proposed by Rotemberg and Woodford (1999) in their analysis of cyclical markup behavior. As discussed in the text, deviations from the baseline include: CES production, overhead labor, marginal wage differing from the average wage due to an overtime premium, and convex costs of adjusting labor.

As will become clear, the CES, overhead labor, and adjustment cost models all alter the marginal product of labor. They accordingly affect the measures of both the overall gap and the price markup, but not the wage markup. On the other hand, the marginal wage model alters only the composition of the gap between the price and wage markups at each point in time, without influencing the gap variable itself.

1. Baseline Specification

Our baseline case assumes no adjustment costs and a production function isoelastic in labor, that is, $Y_t = F(X_t)N_t^{\alpha}$. In this case we have the following expressions for the (log) marginal product of labor and the price markup (up to an additive constant):

$$\begin{aligned} mpn_t &= y_t - n_t, \\ \mu_t^p &= p_t - (w_t - mpn_t) \\ &= -s_t, \end{aligned}$$

where s_t is the log labor share. These two formulas are then used in conjunction with information on the households' marginal rate of substitution to obtain measures of the gap and the wage markup.

2. CES Technology

Here we assume a CES production function: $Y_t = [(1-\alpha)K_t^{1-1/\nu} + \alpha(Z_tN_t)^{1-1/\nu}]^{\nu/(\nu-1)}$. The implied elasticity of output with respect to labor input, $\eta_t \equiv \partial Y_t/\partial N_t/N_t/Y_t$, is given by

$$\eta_t = 1 - (1-\alpha)\left(\frac{Y_t}{K_t}\right)^{-(1-1/\nu)}.$$

Log-linearizing around a steady state yields (ignoring constants) log $\eta_t = \vartheta(y_t - k_t)$, where $\vartheta \equiv (1-\nu^{-1})(\eta^{-1}-1)$. Because $\mathrm{MPN_t} = \eta_t Y_t/N_t$, we can write

$$mpn_t = (y_t - n_t) + \vartheta(y_t - k_t)$$

and

$$\begin{aligned} \mu_t^p &= p_t - (\mathrm{w_t} - \mathrm{mpn_t}) \\ &= -s_t + \vartheta(y_t - k_t). \end{aligned}$$

Calibration of ϑ proceeds by first noticing that the gross price markup M_t^p equals $MPN_t/(W_t/p_t)$. This allows allows us to derive a simple expression for the steady state value of the elasticity of output with respect to labor as a function of the steady state price markup and the labor share, that is,

$\eta = SM^p$. Rotemberg and Woodford (1999) calibrate the coefficient ϑ using approximate values for the average labor income share ($S = 0.7$), the average gross price markup (close to unity), and an estimate for elasticity of substitution between capital and labor ($\nu = 0.5$), all of which combined yield a value $\vartheta = -0.4$.

3. Overhead Labor

For this case we assume a technology given by the production function $Y_t = Z_t K_t^{1-\alpha}(N_t - N_t^*)^\alpha$, where N_t^* denotes the amount of overhead labor at each point in time. The elasticity of output with respect to (total) labor input is now given by

$$\eta_t = \alpha\left(\frac{N_t}{N_t - N_t^*}\right).$$

Log-linearizing around the steady state and ignoring constants yields $\log \eta_t = -\delta \hat{n}_t$, where $\delta \equiv N^*/N - N^*$ is the steady-state ratio of overhead to variable labor, and n_t denotes the log deviation of hours from its long-run trend (around which the linearization is carried out). Using the fact that $MPN_t = \eta_t Y_t / N_t$, it follows that

$$mpn_t = y_t - n_t - \delta \hat{n}_t$$

and

$$\mu_t^p = p_t - (w_t - mpn_t)$$
$$= -s_t - \delta \hat{n}_t.$$

Rotemberg and Woodford (1999) use a zero-profit condition in steady state in order to calibrate δ. In particular, it can be shown that the ratio of average costs to marginal costs can be written as

$$\frac{AC_t}{MC_t} = 1 + \alpha\left(\frac{N_t^*}{N_t - N_t^*}\right).$$

This implies the following steady-state relationship:

$$AC = \frac{1}{M} + \frac{\delta}{1+\delta} S.$$

Following Rotemberg and Woodford (1999), we assume $S = 0.7$, $M = 1.25$, and impose the zero-profit condition $AC = 1$, thus implying $\delta = 0.4$. We use the latter value to construct our overhead-labor measure of the gap and the price markup.

4. Marginal Wage Different from Average Wage

In the previous analysis we have assumed that firms are wage-taking, so that the marginal wage is equal to the average wage. As emphasized by Bils (1987), this will not be the case if the wage rises as firms ask their employees to work more hours. The relevant wage needed to compute both the price and wage markups is no longer the average wage but the marginal wage, W^m. Notice however, that the use of the marginal wage will only alter the decomposition of our gap measure between the price and the wage markup, but not the gap measure itself.

Let $q_t \equiv w_t^m - w_t$ denote the ratio of the marginal to the average wage (expressed in logs). Then it follows that

$$\mu_t^p = p_t - (w_t^m - mpn_t)$$
$$= p_t - (w_t - mpn_t) - q_t$$
$$= -s_t - q_t.$$

Similarly,

$$\mu_t^w = (w_t^m - p_t) - mrs_t$$
$$= (w_t - p_t) - mrs_t + q_t.$$

Bils (1987) and Rotemberg and Woodford (1999) propose a simple model of overtime pay that implies that the ratio Q_t is an increasing function of hours per worker H_t. Log-linearization of that function around a steady-state value for hours per worker allows us to rewrite the price markup as

$$\mu_t^p = p_t - (w_t - mpn_t) - \tau \hat{h}_t,$$

where τ is the elasticity of the marginal-to-average wage ratio with respect to hours per worker.

Similarly, the wage markup will now be given by

$$\mu_t^w = (w_t - p_t) - mrs_t + \tau \hat{h}_t.$$

As discussed in Bils (1987), the assumption of a 50% overtime premium (the statutory premium in the United States) implies $\tau = 1.4$, which we use to construct our overtime measure of the price and wage markups.

5. Labor Adjustment Costs

Finally, we consider the implications of having a cost of adjusting labor, which we assume take the form of output lost. Those cost are to be taken into account when computing firms' marginal costs and hence price markups. Following Rotemberg and Woodford (1999), we assume that those costs take the form $U_t N_t \phi(N_t/N_{t-1})$, where U_t is the price of the input required to make the adjustment. In this case, the (expected) total cost associated with hiring an additional worker for one period is given by

$$W_t\left\{1 + \frac{U_t}{W_t}\left[\phi\left(\frac{N_t}{N_{t-1}}\right) + \frac{N_t}{N_{t-1}}\phi'\left(\frac{N_t}{N_{t-1}}\right) - E_t\left\{R_{t,t+1}\frac{U_{t+1}}{U_t}\left(\frac{N_{t+1}}{N_t}\right)^2\phi'\left(\frac{N_{t+1}}{N_t}\right)\right\}\right]\right\} \equiv W_t B_t,$$

where $R_{t,t+1}$ is the usual stochastic discount factor for one-period-ahead income.

Hence, the expression for the price markup is given by

$$\mu_t^p = p_t - (w_t + b_t - mpn_t)$$
$$= -s_t - b_t.$$

Assuming that the ratio U_t/W_t is stationary, we can derive the following expression in terms of deviations from steady state as follows (ignoring constants):

$$b_t = \xi(\Delta n_t - \beta E_t\{\Delta n_{t+1}\}),$$

where $\xi \equiv (U/W)\phi''(1)$ and $\beta = R\gamma_u$ with γ_u being the steady-state value for U_{t+1}/U_t. Hence, the expression for the price markup can now be written as

$$\mu_t^p = -s_t - \xi(\Delta n_t - \beta E_t\{\Delta n_{t+1}\}).$$

We construct our adjustment-cost measure of price markups under the assumption that $\beta = 0.99$ and $\xi = 4$, the values suggested by Rotemberg and Woodford (1999).