MARKUPS, GAPS, AND THE WELFARE COSTS OF BUSINESS FLUCTUATIONS

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

Banco de España

Markups, Gaps, and the Welfare Costs of Business Fluctuations[*]

Jordi Galí [†]   Mark Gertler[‡]   J.David López-Salido[§]

June 2002
(first draft: May 2001)

Abstract

In this paper we present a simple, theory-based measure of the variations in aggregate economic efficiency associated with business fluctuations. We decompose this indicator, which we refer to as "the gap", into two constituent parts: a price markup and a wage markup, and show that the latter accounts for the bulk of the fluctuations in our gap measure. We also demonstrate the connection between our gap measure and the gap between ouput and its natural level, a more traditional indicator of aggregate inefficiency. Finally, we derive a measure of the welfare costs of business cycles that is directly related to our gap variable. Our welfare measure corresponds to the inefficient component of economic fluctuations, and should thus be interpreted as a lower bound to the costs of the latter. When applied to postwar U.S. data, for some plausible parametrizations, our measure indicates non-negligible welfare losses of gap fluctuations. The results, however, hinge critically on some key parameters, including the intertemporal elasticity of labor supply.

JEL Classification: E3

Key Words: Business Cycles, Countercyclical Markups, Welfare Costs

1

1   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 where 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 develop - 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 zero 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 welfare costs of business cycles.

From a somewhat different perspective, as we discuss below, the inefficiency gap corresponds to the inverse 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 allowing for the possibility that the movement in the overall markup depends on variations in a wage markup as well as in a price markup.[1] Put differently, in contrast to much of the existing literature, we allow for the possibility of labor market frictions that introduce a wedge between the wage and households' consumption/leisure tradeoff. By doing so we can obtain some sense of the relative importance of price versus wage rigidities for overall fluctuations in the inefficiency gap. In addition, focusing on the gap between the labor demand and supply curves leads directly to a measure of the welfare costs of variations in aggregate inefficiency associated with business cycles, based on the lost surplus owing to aggregate fluctuations.

Our approach builds on a stimulating paper by Hall (1997) that analyzes the cyclical behavior of the neoclassical labor market equilibrium. Specifically Hall shows 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. Hall interprets this difference – which we refer to as the Hall residual – as reflecting a preference shock[2]. However, we present evidence that suggests that this residual may instead reflect countercyclical markup

---

[1]See Rotemberg and Woodford (1999) for a survey of the literature on business cycles and countercyclical markups. For business cycle models that feature a role for wage markups as well as price markups, see Blanchard and Kiyotaki (1987) and Erceg, Henderson and Levin (2000).

[2]To be fair, Hall does not take the preference shock interpretation literally, but rather as a starting point for analyzing the significance of the cyclical movement in the gap between the measured labor demand and supply curves. In his conclusion, he observes that cyclical movements in unemployment could underlie the measured preference shock.

2

variation. As we show, under this interpretation, cyclical variation in this residual reflects efficiency costs.

In related work, Mulligan, (2002) also develops a measure of aggregate inefficiency based on the labor market distortion, and examines its long term behavior, using annual data spanning more than a century. He finds that marginal tax rates correlate well at low frequencies with his distortion measure. We instead focus on fluctuations in the degree of aggregate inefficiency at business cycle frequencies, and stress countercyclical markup variation (in a very broad sense) as the key factor underlying fluctuations at these frequencies. We also show how the labor market distortion at the business cycle frequency is related to the gap between output and its natural level, a more conventional indicator of cyclical inefficiency. Finally, we describe how to construct a welfare cost measure based on the observed labor market distortion.

In section 2 we develop a framework for measuring the inefficiency gap and its price and wage markup components in terms of observables, conditional on standard assumptions about preferences and technology. In section 3 we present empirical measures of this variable for the postwar U.S. economy. We show that the inefficiency gap exhibits large procyclical swings, thus confirming our basic hypothesis. In addition, most of its variation is associated with countercyclical movements in the wage markup.[3] The price markup shows, at best, a weak contemporaneous correlation. Finally, we demonstrate the robustness of our gap measure and its decomposition to alternative assumptions about preferences and technology. In section 4 we consider the possibility that preference shocks underlie the variation in our gap measures. Specifically, we present VAR evidence that suggests that the Hall residual 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. Section 5 characterizes both theoretically and empirically the link between the labor market distortion and the output gap.

In Section 6 we examine the welfare consequences of business fluctuations. Our approach differs significantly from Lucas (1987), who examines the welfare costs of consumption variability associated with the cycle. We instead focus on the welfare costs associated with fluctuations in the efficiency of resource allocation, as implied by the time series variation in our gap measure. As we show, our framework implies that business contractions below the steady state produce greater efficiency costs than the efficiency gains arising from symmetric expansions above. This asymmetry, in turn, implies that fluctuations raise efficiency costs on average, even if these fluctuations are themselves symmetric. We show that under some not implausible parametrizations, gap fluctuations of the magnitude observed in postwar U.S. data could potentially

---

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

involve non-negligible welfare costs. The results, however, depend critically on several key parameters, including the intertemporal elasticity of labor supply. Finally, in addition to studying average performance over the postwar, we also examine the efficiency costs associated with the major boom-bust episodes of the 1970s and early 1980s. Concluding remarks are in section 7.

2 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 mpn_t and mrs_t denote, respectively, the (log) marginal product of labor and (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 $\{gap_t\}$ 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 non-stationary, as it 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 unit of labor input (including non-wage costs).[4] Accordingly, we define the aggregate price markup as follows:

$$
\begin{aligned}
\mu_t^p &= p_t - (w_t - mpn_t) \tag{2} \\
&= mpn_t - (w_t - p_t) \tag{3}
\end{aligned}
$$

The aggregate wage markup is given by:

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

i.e., 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.

---

[4]We show subsequently that our results are robust to allowing for labor adjustment costs.

4

There are a variety of frictions (perhaps most prominently, wage and price rigidities) which 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.[5] Nominal price rigidity, in turn, may give rise to a countercyclical price markup in response to demand shocks since, holding productivity constant, the marginal product of labor is countercyclical.[6]

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 (i.e., the gap is always negative), so that (small) increases in our gap measure will be associated with a smaller distortion (i.e., 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. Identifying the markups requires some assumptions about technology and preferences. We first consider a baseline case with reasonably conventional assumptions. We then show that our results are robust to a number of leading alternative restrictions.

Given equation (2), identification of price markup variations only requires an assumption on technology. Under the assumption of a technology with constant elasticity of output with respect to hours (say, α), we have (up to an additive constant):

---

[5]Models with countercyclical wage markups due to nominal rigidities include Blanchard and Kiyotaki (1987) and Erceg, Henderson and Levin (2000). Alexopolous (2000) develops a model with a real rigidity due to effiiciency wages that can generate a countercyclical wage markup.

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

<div align="center">5</div>

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

where y_t is output per capita and n_t is hours per capita.[7]

Combining equations (2)and (8) yields:

$$\mu_t^p = (y_t - n_t) - (w_t - p_t) \tag{9}$$
$$\equiv -ulc_t \tag{10}$$

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

Let c_t be consumption per capita and $\overline{\xi}_t$ be a deterministic, low frequency preference shifter. Then, the (log) marginal rate of substitution can be written (up to an additive constant) as:

$$mrs_t = \sigma\ c_t + \varphi\ n_t - \overline{\xi}_t \tag{11}$$

where the parameter σ is the coefficient of relative risk aversion and φ measures the curvature of the disutility of labor. Following Hall (1997), we allow for the possibility of low frequency shifts in preferences over consumption versus leisure, as represented by movements in $\overline{\xi}_t$. 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 4 we provide evidence to justify this assumption. It follows that the wage markup is given by:

$$\mu_t^w = (w_t - p_t) - (\sigma\ c_t + \varphi\ n_t) + \overline{\xi}_t \tag{12}$$

Given a measure of both the price and the wage markup, one can obtain a measure of the gap using equation (6). Alternatively, one can combine equations (9), (12) and (6) to obtain:

$$gap_t = (\sigma\ c_t + \varphi\ n_t - \overline{\xi}_t) - (y_t - n_t) \tag{13}$$

3 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

---

[7]As we discuss in section 5, this specification of production allows for variable capital utilization. Under certain conditions it is also compatible with variable labor utilization, particularly if labor effort moves roughly proportionately with hours, as the evidence suggests (see, e.g., Basu and Kimball, 1997).

6

is based on quarterly postwar U.S. data over the sample period 1960:4 - 1999:4, and are drawn from the DRI database. The basic data used to construct the gap variable and its components include compensation per hour (LBCPU), hours all persons (LBMNU), real and nominal output (GPBUQ and GPBU), all of which refer to the refer to the nonfarm business sector, as well as the NIPA series for non-durable and services consumption (GCNQ+GCSQ). In addition we also use population over sixteen (P16) (to express variables in per capita terms), real GDP (GDPQ), implicit GDP deflator (GDPD), the Fed-funds rate (FYFF), the spread between the 10-year government bond yield (FYGL) and the 3-month Treasury Bill rate (FYGM3), and a commodity price index (PSCOM) for our VAR exercise in Figure 7.

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 which corresponds to the inverse of the (Frisch) wage elasticity of labor supply. A vast amount of evidence from micro-data suggests wage elasticities mostly concentrated in the range of $0.05 - 0.3$.[8] The business cycle literature tends to use much higher values. We accordingly use a baseline value $\varphi = 5$, which corresponds to a labor supply elasticity of 0.2, which is slightly above the mean of the labor supply elasticity estimates from the micro data.[9] However, we also experiment with other values, including values used in the business cycle literature (see, e.g., Cooley and Prescott, 1995.)

There is a similar controversy over the choice of the coefficient of relative risk aversion, which corresponds to the inverse of the intertemporal elasticity of substitu-

---

[8]MacCurdy (1981) estimates the Frisch elasticity of labor supply for men to be 0.15, a finding that has been largely confirmed by subsequnet literature (e.g., Altonji,1986, and more recently Pencavel, (forthcoming)). In his survey of the literature, Card (1994) concludes that this elasticity is "surely no higher than 0.5 and probably no higher than 0.2." See Mulligan (1998), however, for an alternative view. Finally, less is kown about female labor supply elasticity, but Pencavel (1998) has estimated a Frisch elasticity of 0.21 for this group. Pencavel's sample period covers the mid 1970s to mid 1990s. Apparently, as the gap betweeen male and female labor force participation has narrowed, female labor supply elasticities have become similar to male labor supply elasticities.

[9]Whether it is appropriate to use the existing micro evidence to calibrate the intertemporal elasticity of labor supply in a business cycle model is a matter of controversy, particularly to the extent employment adjusts along the extensive margin as well as the intensive margin (see, e.g. the discussion in Mulligan (1998)). How to take into account the extensive margin as well as other pertinent frictions relevant to intertemporal labor supply decisions is a tough issue. In our view it is best to just think of φ as a way to parametrize in a reduced form sense how the opportunity cost of labor changes over the business cycle. A value of $\varphi = 5$, for example, implies that everything else equal, a one percent reduction in aggregate hours worked leads to a five percent reduction in the opportunity cost of labor (since φ correspond to the elasticity of disutility of hours.) For example, a 3.0% reduction in employment, as occurs in a typical recession, would imply that the opportunity cost of labor is worth eighty-five percent of its steady state value. To the extent there are frictions that make unemployment less desirable than the frictionless model implies, this kind a number may not be unreasonable. While for internal consistency we appeal to the micro literature to calibrate the model, we think the best strategy at this stage is to explore the sensitivity of the results to different parameter values.

7

tion. 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 3.3.[10] The business cycle literature instead tends to assume log utility over consumption (i.e., $\sigma = 1$), based on the justification that these preferences are consistent with balanced growth. We also adopt this assumption for our baseline calibration, but then also experiment with parametrizations consistent with the direct evidence.

In addition, we need to make an assumption to identify the low frequency shifter $\overline{\xi}_t$. Let $\widetilde{\mu}_t^w \equiv (w_t - p_t) - (\sigma c_t + \varphi \; n_t)$ be the observable component of the wage markup. It follows that

$$\widetilde{\mu}_t^w = \mu_t^w - \overline{\xi}_t \tag{14}$$

From this perspective, the wage markup μ_t^w is the cyclical component of $\widetilde{\mu}_t^w$ and $\overline{\xi}_t$ is (minus) the trend component. Following Hall (1997) we approximate the low frequency movements of $\widetilde{\mu}_t^w$ by fitting a fifth-order polynomial of time to $\widetilde{\mu}_t^w$.[11]

Finally, before proceeding, we note that the relationships derived in the previous section hold only up to an additive constant. Accordingly, our framework only allows us to identify the *variations* over time in the markup and its components, but not their levels. Our baseline results thus employ measures of the price and wage markups and the gap constructed using, respectively, equations (10), (12), and (6), and expressed in terms of deviations from their respective sample means.

Figure 2 presents the times series measure of our gap variable under our baseline assumptions of $\sigma = 1$ and $\varphi = 5$. 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). It is also interesting to observe that the gap hovers near zero for most of the period post 1995. The resulting implication is that the rapid output growth over this period must have been due to real factors (e.g. technology improvements) as opposed to excess demand.

We next decompose the movements of the gap into its wage and price markup components. The wage markup measures were constructed using (12).[12] The price markup corresponds to minus the log of real unit labor costs, as implied by (10). Figure 3 shows the behavior of the gap against the wage markup. To facilitate visual inspection, we plot the inverse of the wage markup (i.e., minus the log wage markup). By definition, the difference between the gap and the inverse wage markup

---

[10] Using micro-data, Barsky *et. al* (1997) estimate an intertemporal elasticity of substitution of 0.18, implying a coefficient of relative risk aversion slightly above 5. Using macro-data, 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 polynominal instead of a band pass filter to detrend the data. The fifth order polynominal detrended, however, produces cycles that closely resemble those that arise from a band pass filter that removes frequencies above 2 quarters and below either 32 or 64 quarters,

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

8

is the inverse price markup. What is striking about the pictures is the strong co-movement between the gap and the (inverse) wage markup. Put differently, our evidence suggests that the inefficiency gap seems is driven largely by countercyclical movements in the wage markup.[13]

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 hence that they can be used to construct a relevant cost measure.[14] While this assumption is standard in the literature on business cycles and markups (e.g., Rotemberg and Woodford, 1999), it is not without controversy. Notice, however, that even if observed wages were not allocational, our gap variable would still be correct (since its construction does not require the use of wage data), though its decomposition between wage and price markups would most likely be distorted by the use of an incorrect measure of the true (shadow) cost of labor.

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 markup, and also for detrended (log) GDP, a common indicator of the business cycle. Note first that the percent 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].

In Figure 4 we demonstrate that the qualitative pattern of the gap is robust to reasonable alternative assumptions about the coefficient of relative risk aversion and the labor supply elasticity. We first explore adjusting the coefficient of relative risk aversion. We consider a value of 5 for this parameter, implying an intertemporal elasticity of substitution of 0.2, consistent with the evidence mentioned earlier. The top panel of Figure 4 plots our gap variable for the case of $\sigma = 5$ versus the baseline case of $\sigma = 1$ (In each instance we keep φ at its baseline value of 5). Clearly, the qualitative pattern is similar across the cases. The amplitude of the gap variable, though,

---

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

[14]Some indirect evidence that wages are allocational is found in Sbordone (1999) and Gali 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 we showed in an earlier version of this paper, 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 co-movement 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 (1999) and Gali and Gertler (1999).

9

increases with risk aversion. Intuitively, a rise in risk aversion makes labor supply more inelastic, which raises the sensitivity of the gap to employment fluctuations.

While the micro-evidence suggests a small labor supply elasticity, the business cycle literature tends to assume a high elasticity, typically unity and above. We accordingly reconstruct the gap measure assuming $\varphi = 1$, which implies a Frisch labor supply elasticity of unity. The bottom panel of Figure 4 plots the behavior of the gap under this new parametrization against the baseline case of $\varphi = 5$. Again, the qualitative pattern is similar across the two cases. A higher supply elasticity, however, does imply quantitatively smaller fluctuations in the gap. Intuitively, a more elastic labor supply curve implies that any change in employment from its natural level will yield a smaller change than otherwise in the distance between the labor demand and labor supply curves.

Though we do not report the results here, it also remains true that the movements in the gap for both the high labor supply elasticity case and the high risk aversion case are associated largely with countercyclical movements in the wage markup. This should not be surprising since the wage markup is computed simply as minus the difference between the gap and the price markup, where the measure of the latter is invariant to the labor supply elasticity.

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

Finally, we demonstrate that our gap measure is robust to alternative assumptions about production (that yield alternative measures of the marginal product of labor.) Our baseline case assumes constant elasticity of output with respect to hours. We consider three alternative assumptions suggested by Rotemberg and Woodford (1999): (i) Cobb-Douglas modified to allow for overhead labor; (ii) CES; and (iii) Cobb-Douglas with labor adjustment costs. In each case we follow the parametrization recommended in Rotemberg and Woodford (1999). As Figure 5 indicates, our gap measures are quite robust to these alternative assumptions. Though we do not report the results here, it remains the case that the movements in the gap are strongly associated with a countercyclical wage markup. For completeness, Figure 6 shows that our results are robust to allowing for a measure of the marginal rate of substitution based on time dependent preferences in leisure, following Eichenbaum, Hansen and Singleton (1988).

10

4   Preference Shocks, Hall's Residual and the Gap

As we have discussed, the notion that the business cycle is associated with cyclical movements in the gap between the labor demand curve and the observable component of the labor supply curve originated with Hall (1997). In his baseline identification scheme, however, Hall associated this gap entirely with preference shocks.[16] In this section we show that the high frequency movements in the gap cannot be simply due to preference shocks.

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

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

with

$$\xi_t = \overline{\xi_t} + \widetilde{\xi}_t$$

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 of the marginal rate of substitution allowing for preference shocks (15), together with (8) and the definition of the inefficiency gap (1), it is possible to express Hall's residual as follows:

$$\begin{align} x_t &\equiv (mrs_t - mpn_t) + \xi_t \tag{17} \\ &= -(\mu_t^p + \mu_t^w) + \xi_t \tag{18} \end{align}$$

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 variable x_t as a preference shock, since under this assumption $x_t = \xi_t$.[17] Notice that under these circumstances the efficiency gap is zero, 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 for

---

[16]As we noted earlier, Hall does not take the preference shock hypothesis literally. However, since it is a possible alternative to our countercyclical markup interpretation, it is worth investigating this possibility.

[17]See also Baxter and King (1991). Holland and Scott (1998) construct similar measures for the U.K.

11

departures from perfect competition, x_t will purely reflect movements in markups, i.e., $x_t = -(\mu_t^p + \mu_t^w)$. In the latter instance, Hall's residual corresponds exactly to our inefficiency gap, i.e.,, $x_t = gap_t$, for all t.

Note that if the Hall residual 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 the Hall residual should be exogenous. We next present 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 a rough measure of the stance of monetary policy, while detrended GDP is just a simple cyclical indicator. Table 2 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. This finding is robust to reasonable alternative parametrizations of σ and φ. Overall, the evidence of Granger causality is inconsistent with the hypothesis that the Hall residual 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 et al. (1999), and others. It is based on a VAR that includes measures of output, the price level, commodity prices, and the Federal 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 Funds rate, under the assumption that this shock does not have a contemporaneous effect on the other variables in the system. Figure 7 shows the estimated responses to a monetary contractionary policy shock. The responses of the nominal rate, output 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 respect, note that the tight money shock induces a rise in the wage markup that closely mirrors the decline in the gap, both in the shape and the magnitude of the response. This countercyclical movement in the wage markup is consistent with evidence on unconditional variation presented in Table 1. The price markup also rises, though with a significant lag. Apparently, the sluggish response of wages, which gives rise to strong countercyclical movement in the wage markup,

12

delays the rise in the price markup.[18] 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. While 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, as seems plausible under the interpretation that they largely reflect institutional and demographic factors, then they are likely to be captured by the trend component associated with our low frequency filter. In this instance our filtered gap series, which isolates the high frequency movement in this variables, is likely to be largely uncontaminated by exogenous preference variations.

5   Relation to the Output Gap

In this section we illustrate the connection between our gap measure and the output gap, a more traditional indicator of cyclical utilization. The output gap is commonly meant to refer to the deviation of output from its *natural* level, defined as the equilibrium value in the absence of nominal rigidities. Formally,

$$\widetilde{y}_t \equiv y_t - \overline{y}_t \tag{19}$$

where $\widetilde{y}_t$, and $\overline{y}_t$ denote the output gap and the natural level of output, respectively. While it is not possible to derive an exact relation between the output gap and the inefficiency gap without specifying a complete model, we demonstrate how it is possible to derive a relatively tight band simply conditional on a minimal set of plausible assumptions.

First, we need a restriction on technology. For simplicity, we assume that the reduced form aggregate production function can be written as:

$$y_t = a \; n_t + z_t \tag{20}$$

where z_t is exogenous or, at least, invariant to the degree of nominal rigidities.[19] Think of z_t as including both technology and capital, where we treat capital as exogenous on the grounds that the percent fluctuations in capital at the business cycle

---

[18]As Gali and Gertler (1999) and Sbordone (1999) 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 inverse of the price markup. The sluggish response to the price markup translates into sluggish behavior of real marginal cost.

[19]Simulations using a model with capital accumulation suggests that such an assumption is a good approximation.

frequency a relatively small. Note that equation (20) allows the possibility of variable capital utilization, following Burnside and Eichenbaum (1996) and King and Rebelo (1999). Here we simply observe that variable capital utilization will raise the *effective* output elasticity of employment, a. Given equation (20), we can derive the following expression for the gap:

$$gap_t = \left(\frac{1-a+\varphi}{a}\right) y_t + \sigma c_t - \left(\frac{1+\varphi}{a}\right) z_t - \overline{\xi}_t \qquad (21)$$

Second, we assume that the only source of gap variation lies in the presence of nominal rigidities in labor and/or goods markets. In other words, frictionless or desired markups are assumed to be constant. This assumption permits us to interpret the natural level of output as the level of output consistent with a *constant* gap (which corresponds to its steady state value, gap). Accordingly, if we let $\overline{c}_t$ be the level of consumption in the absence of nominal rigidities, then it follows that $\overline{y}_t$ satisfies

$$gap = \left(\frac{1-a+\varphi}{a}\right) \overline{y}_t + \sigma \overline{c}_t - \left(\frac{1+\varphi}{a}\right) z_t - \overline{\xi}_t \qquad (22)$$

To obtain a relation between the output gap our demeaned gap measure $\widehat{gap}_t \equiv gap_t - gap$ (e.g., as portrayed in figure 2), first combine equations (21) and (22):

$$\widehat{gap}_t = \left(\frac{1-a+\varphi}{a}\right) \widetilde{y}_t + \sigma \widetilde{c}_t \qquad (23)$$

where $\widetilde{c}_t = c_t - \overline{c}_t$. Without loss of generality, we can express the consumption gap, $\widetilde{c}_t$, as a time varying proportion of the output gap, as follows

$$\widetilde{c}_t = \eta_t \, \widetilde{y}_t \qquad (24)$$

By establishing a reasonable band for η_t, we can derive a band for $(y_t - \overline{y}_t)$ as a function of the our gap measure. To see this, notice that by substituting equation (24) into (23), we obtain,

$$\widetilde{y}_t = \left(\frac{a}{1+\varphi+a(\sigma\eta_t-1)}\right) \widehat{gap}_t \qquad (25)$$

which allows us to derive upper and lower bounds for the output gap, for any give bounds on η_t.

We illustrate the previous approach by computing the output gap for our baseline preference specification, with $\sigma = 1$ and $\varphi = 5$. We follow King and Rebelo (1999), who argue that the evidence is consistent with a value of a of roughly unity. Finally, we consider a range of values for η_t in the interval 0.6 to 1.0. A reasonable upper bound for η_t is unity, given that consumption is smooth relative to the durable components of output. A reasonable lower bound is a fraction in the vicinity of 0.6, given that the

14

standard deviation of consumption is roughly 0.8 the standard deviation of output, with a correlation of roughly 0.9. (See, e.g., Campbell and Deaton, 1989.)

Figure 8 plots the output gap for both the upper and lower bounds on η_t. For comparison, it also plots a commonly used measure of the output gap, constructed by the Congressional Budget Office (CBO). Broadly speaking, our measure of the output gap has similar properties to the CBO gap, the sample correlation between both series being 0.74. In addition, the bands on our output gap measure turn out to be very tight, suggesting that our estimate of that variable is relatively insensitive to reasonable variation in η_t. These results appear to be largely robust to reasonable choices of the labor supply elasticity parameter.

6 Welfare and the Gap

In this section we derive a simple way to measure the welfare costs of business fluctuations associated with variations in our gap variable and then apply this methodology to postwar U.S. data. In addition to obtaining a measure of the average cost of gap fluctuations, we also apply our methodology to examining the efficiency losses during particular episodes, including the major postwar recessions.

As we noted in the introduction, our approach differs from Lucas (1987) who considered the costs to risk averse households of the consumption variability associated with the cycle. For roughly the same reason that the baseline neoclassical model has difficulty accounting for the equity premium (i.e., the relatively low variability in aggregate consumption), the Lucas approach suggests very low costs of business fluctuations. For reasonable degrees of risk aversion, Lucas finds that households would be willing to sacrifice less than 0.1 percent of their consumption per period to eliminate fluctuations, clearly a small number. Many papers have extended the Lucas approach, either to allow for incomplete markets or to allow the business cycle to have more persistent effects on consumption variability. These papers also tend to find small welfare costs, though with a few exceptions (e.g., Barlevy (2000) and Beaudry and Pages (2001)).

Our approach instead measures the costs stemming from fluctuations in the degree of inefficiency of the aggregate resource allocation, as reflected by the movements in our gap variable.[20] As in Ball and Romer (1987), the cycle generates losses on average within our framework 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 steady state price and wage markups), 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

---

[20]For a very early attempt to measure the welfare cost of inefficiently high unemployment, see Gordon (1973).

15

the steady state and falls when employment moves 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. Below we show that under some plausible parameter values the resulting net welfare cost can be substantial.

6.1   A Welfare Measure

We now proceed to derive our welfare measure. We continue to assume a production technology with constant elasticity of output with respect to hours, as given by equation (20). We also assume, following King and Rebelo, that the output elasticity with respect to hours is unity ($a = 1$). Accordingly, equation (20) can be written (in levels) as

$$Y_t = Z_t \; N_t \tag{26}$$

Next, we obtain a measure of the utility gain or loss, Δ_t, from reallocating employment at time t from its existing level N_t to the level that would arise in the frictionless equilibrium, $\overline{N}_t$. Let $W(N_t)$ be the period t utility value of output, net of the utility cost of working, conditional on employment level N_t. Accordingly,

$$\Delta_t = W(N_t) - W(\overline{N}_t) \tag{27}$$

Assuming that, at the margin, the household is indifferent between consuming and saving any additional unit of output, the shadow value of the latter will equal the marginal utility of consumption. Accordingly, if $U(C_t, N_t)$ is the period utility function of the representative consumer, then the gross utility gain from raising employment is given by the marginal utility of consumption times the marginal product of labor, i.e. $U_{c,t} \; Z_t$. The cost is $U_{n,t} < 0$, the marginal disutility of hours. On net the welfare effect thus given by $W'(N_t) = U_{c,t} \; Z_t + U_{n,t}$.

Ultimately, however, we are interested in an expression for the total utility gap, Δ_t. If the percent difference between N_t and $\overline{N}_t$ is not large, then it is reasonable to approximate $W(N_t)$ with a second order Taylor expansion about $W(\overline{N}_t)$. Accordingly, assuming that utility is separable in consumption and leisure (i.e., $U_{cn} = 0$) we have:

$$\begin{aligned}
\Delta_t \; &\simeq \; [\overline{U}_{c,t} Z_t + \overline{U}_{n,t}] \; (N_t - \overline{N}_t) \\
&+ \frac{1}{2} \left(\overline{U}_{cc,t} \; \frac{\partial \overline{C}_t}{\partial \overline{Y}_t} \; Z_t + \overline{U}_{nn,t} \right) \; (N_t - \overline{N}_t)^2
\end{aligned} \tag{28}$$

where, in order to lighten the notation we have defined $U_t \equiv U(C_t, N_t)$, $\overline{U}_t \equiv U(\overline{C}_t, \overline{N}_t)$, and $U \equiv U(C, N)$.

Assume that preferences are given by $U(C, N) = \frac{C^{1-\sigma}}{1-\sigma} - \frac{N^{1+\varphi}}{1+\varphi}$. Then, as we show in the appendix, it is possible to express Δ_t as the following quadratic function of

16

the percent deviation of employment from its value in the frictionless equilibrium, $\widetilde{n}_t \equiv \log\left(\frac{N_t}{\overline{N}_t}\right)$:

$$\Delta_t \simeq \frac{\overline{U}_{c,t}\overline{Y}_t}{(1+\mu)} \left\{ \mu \, \widetilde{n}_t \, - \frac{1}{2} \left[\varphi + \sigma(1+\mu)\overline{\eta}_{cy,t} - \mu \right] \, \widetilde{n}_t^2 \right\} \tag{29}$$

where $\overline{\eta}_{cy,t} = \frac{\partial \overline{C}_t}{\partial \overline{Y}_t} \frac{\overline{Y}_t}{\overline{C}_t}$ is the elasticity of consumption with respect to output in the frictionless equilibrium and.where, as before, μ is the steady state net markup.

Since we do not have a direct measure of the frictionless level of employment, $\widetilde{n}_t$ is not observable. We can, however, derive a relation between $\widetilde{n}_t$ and our gap variable, $\widehat{gap}_t$, which is measurable. Given the production function (26), it follows that $\widetilde{n}_t = \widetilde{y}_t$. We can then obtain a relation for $\widetilde{n}$ in terms of $\widehat{gap}_t$, by exploiting the analysis of the previous section, in particular equation (25). For simplicity, we assume $\overline{\eta}_{cy,t} = 1$, but as in the previous section, our results are not sensitive to reasonable variations in the parameter. Given this restriction and our earlier assumption that $a = 1$, it follows that

$$\widetilde{n}_t = \left(\frac{1}{\sigma + \varphi}\right) \, \widehat{gap}_t \tag{30}$$

Combining equations (29) and (30) then yields an expression for Δ_t in terms of $\widehat{gap}_t$:

$$\Delta_t \simeq \overline{U}_{c,t}\overline{Y}_t \, \omega(\widehat{gap}_t) \tag{31}$$

where

$$\omega(\widehat{gap}_t) = \frac{1}{(1+\mu)(\sigma+\varphi)} \left[\mu \, \widehat{gap}_t - \frac{1}{2}\left(1 + \frac{\mu(\sigma-1)}{\sigma+\varphi}\right) \, \widehat{gap}_t^2 \right] \tag{32}$$

Observe that $\omega(\widehat{gap}_t)$ is the efficiency loss or gain from gap deviations from its steady value, expressed as a percent of the frictionless level of output $\overline{Y}_t$ (since $\omega(\widehat{gap}_t) = \Delta_t/(\overline{U}_{c,t}\overline{Y}_t)$). The first term in brackets, the linear term, reflects the symmetric costs and benefits from the gap moving below and above the steady state, due to the positive steady state markup μ. The second term, the quadratic term, captures the asymmetric effect of gap fluctuations on welfare. In particular, a reduction in the gap below the steady state results in an efficiency loss that exceeds the gain stemming from a commensurate increase in the gap above the steady state. Under the (weak) assumption that $\frac{\mu(\sigma-1)}{\sigma+\varphi} > -1$, ω is a concave function of the gap, implying that the welfare losses from gap contractions are less than made up for by the welfare gains from symmetric gap expansions.

Finally, in keeping with the literature, we express the welfare losses from gap fluctuations measure terms of an equivalent loss in consumption. In particular, we divide Δ_t by $\overline{U}_{c,t}\overline{C}_t$ to make the metric the percent of consumption in the frictionless

17

equilibrium:. Let $\Phi = Y/C$ be the steady state output/consumption ratio. Then the efficiency loss (or gain) as a percent of $\overline{C}_t$ is given by

$$
\begin{aligned}
\frac{\Delta_t}{\overline{U}_{c,t}\overline{C}_t} &\approx \frac{\overline{Y}_t}{\overline{C}_t}\,\omega(\widehat{gap}_t) \\
&\approx \Phi\,\omega(\widehat{gap}_t)
\end{aligned} \tag{33}
$$

where the second approximation holds under the assumption that $\overline{\eta}_{cy,t} \simeq 1$.

We can use equation (33) 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, we take the unconditional expectation of equation (33) to obtain:

$$
E\left\{\frac{\Delta_t}{\overline{U}_{c,t}\overline{C}_t}\right\} = -\frac{\Phi}{(1+\mu)(\sigma+\varphi)}\left[\frac{1}{2}\left(1+\frac{\mu(\sigma-1)}{\sigma+\varphi}\right)\,var(gap_t)\right] \tag{34}
$$

where $var(gap_t)$ is the variance of the inefficiency gap. Notice that, as a result of the concavity of ω, the expected welfare effects of fluctuations in the gap variable are negative, i.e. 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 $var(gap_t)$. As section 3 suggests, $var(gap_t)$.is potentially large if labor supply is relatively inelastic or risk aversion is relatively 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. To see that, suppose 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, while some losses would still be implied by the variability of consumption and leisure (under standard convexity assumptions on preferences).

6.2   Some Numbers

Equation (33) provides a real time measure of the efficiency costs of the cycle, $\Delta_t/\overline{U}_{c,t}\overline{C}_t$, conditional on our gap variable, $\widehat{gap}_t$. Accordingly, we construct a quarterly time series of $\Delta_t/\overline{U}_{c,t}\overline{C}_t$, taking as input our baseline measure of the gap, based on $\sigma = 1$ and $\varphi = 5$. Figure 9 plots the resulting time series over the sample 1960:IV-99:IV. The value at each period t is interpretable as the efficiency gain or loss in percentage units of consumption associated with the deviation of the inefficiency gap from its steady state. The asymmetric nature of the gains and losses is clear. As the figure shows, significant efficiency losses arise in recessions that do not appear to be offset by commensurate gains during booms. Note also that the efficiency losses

18

are largest during the major recessions, ranging between 3.0 and 3.5 percent of consumption per period around the time of the respective troughs. During the major recessions, further, these losses persist for for several years, suggesting nontrivial costs around these periods. We return subsequently to this issue.

We next present a measure of the average welfare cost of the cycle, based on equation (34). As we noted earlier, the measure is simply proportionate to the square of the gap. We construct estimates for alternative values of the parameters φ, σ, and μ. Recall that φ corresponds to the inverse of the Frisch elasticity of labor supply, whereas σ is the inverse of the intertemporal elasticity of substitution. We consider three values of each of these parameters: $1, 5$, and 10, implying that each of the corresponding elasticities ranges from 1.0 to 0.1. For the parameter μ, the sum of the steady state wage and price markups, we consider values of 0.10, 0.25, and 0.40, which we think of as falling within a plausible range.

For a parametrization that corresponds to our baseline case of section 3 augmented with the intermediate value of the markup ($\sigma = 1$; $\varphi = 5$; $\mu = 0.25$), we estimate the welfare cost of postwar U.S. business fluctuations to be roughly 0.26 percent of consumption, overall a modest number, though one that is about three times the size of Lucas' estimate of less than 0.1 percent.

Overall, the estimated welfare costs are highly sensitive to the labor supply elasticity and the intertemporal elasticities of substitution. For high values of the two elasticities (corresponding in our case to $\sigma = 1$ and $\varphi = 1$), the welfare costs are small, on the order of Lucas' estimates. In this case, roughly speaking, the labor supply curve is relatively flat, implying small cyclical fluctuations in the inefficiency gap (see Figure 4.)

On the other hand, in the case of low elasticities, the costs can become fairly significant. With $\varphi = 10$, implying a labor supply elasticity of 0.1, the costs vary from of 0.6 percent of consumption for the case of $\sigma = 1$, to roughly 0.9 for the case of $\sigma = 10$. In this latter case, the labor supply curve is very steep, implying very large fluctuations in the inefficiency gap. It is worth noting that, while higher than normally used in business cycle calibration exercises, values of $\varphi = 10$ and $\sigma = 10$ fall within the range of estimates in the literature, as we discussed in section 3.

Any measure of the average cost of business cycles 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 relatively 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 downturns. Doing so provides a sense of the gains from avoiding future recessions (either by good policy or by good luck.)

We accordingly consider the two major episodes in our sample where the economy experienced a boom followed by a deep recession; the periods 1972-77 and 1978-83.[21]

---

[21] We combine the 1980 and 1981-82 recessions into a single episode.

In each instance we measure the boom as the period where the gap variable climbs above zero up to the point where it returns to zero. The recession is the period that follows, where the gap turns negative up to the time it returns to steady state. For each episode, Table 4 reports the efficiency gains from the boom and the costs from the recession, followed by the net loss (the sum of the first two columns.)

We measure the gains and losses as a percent of one year's consumption. In the first row of table 4, we consider the implications of our baseline case, with $\sigma = 1$ and $\varphi = 5$. In each episode, the costs of the recession outweigh the benefits of the boom, as we might expect, given the asymmetric nature of the cyclical efficiency gains and losses. In addition, the net cost of the cycle (last column) as well as the gross cost of the recession (2nd column) is non-negligible. For example, the gross efficiency loss of the 1974-75 recession was equivalent to 5.7 percent of one year's consumption, while the net cost after deducting the gains from the preceding boom was roughly 4.9 percent. For the 1980-82 recession, the gross and net costs were 4.3 percent and 3.3 percent, respectively.

To get a sense of the range of estimates our parametrizations imply, the next two rows consider the polar extremes of high and low intertemporal elasticities, i.e., $(\sigma = 1, \varphi = 1)$ versus $(\sigma = 10, \varphi = 10)$. In the high elasticity case, the net costs are modest: about 2.5 percent of one year's consumption for the 1974-75 recession, and about 2.2% for the 1980-82 downturn. In the low elasticity case, however, they become significant: roughly 6.0 percent in 1974-75 and 8.5 percent in the 1980-82

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 inflation associated with the economy moving above steady state output. For this reason, our metric may overstate the gains from booms. 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.

7   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 substitution in an environment of frictionless markets. The second instead appeals to countercyclical markups owing to particular market frictions. Any model that produces a countercyclical movement in the wedge between the marginal product of labor and the household's marginal rate

20

of substitution potentially falls into this class. Perhaps a central message of this paper is that the issue of the welfare costs of business cycles cannot be cleanly separated from the issue of which of these business cycle paradigms provides a better description of actual economic fluctuations. We find that with high degrees of intertemporal substitution, the costs of business fluctuations are relatively small, which perhaps should not be surprising since labor supply curves are relatively flat in this setting. On the other hand, with low substitution elasticities (implying that strongly countercyclical markups are needed to explain the data), we find significant welfare costs associated with cyclical fluctuations in the inefficiency of resource allocation. To be sure, the appropriate parametrization of these intertemporal elasticities remains an open question. For the time being, though, we note that there is a considerable body of evidence consistent with the low intertemporal elasticities that we stressed in our analysis, as discussed in section 3.

Finally, we emphasize that our estimates of the efficiency costs of business fluctuations are likely to be conservative because they do not take into account the welfare costs of inflation variability that may be associated with cyclical fluctuations. Recent work by Woodford (1999) and others suggests that these efficiency costs may be highly significant. Accounting for this factor in our overall welfare measure is something we plan for future research.

21

Table 1. Basic Statistics: 1960-1999

Baseline Calibration ($\sigma = 1$, $\varphi = 5$)

Variable	*s.d.*(%)	ρ	Correlation			
			GDP	Gap	Price Mkup	Wage Mkup
GDP	2.7	0.91	1			
Gap	13.6	0.94	0.86	1		
Price Markup	4.3	0.96	0.07	-0.05	1	
Wage Markup	14.1	0.95	-0.85	-0.95	-0.26	1

22

Table 2. Granger Causality Tests (1960-1999)

Baseline Calibration ($\sigma = 1$, $\varphi = 5$)
Bivariate VAR

Variable	4-lags	5-lags	6-lags
GDP	0.000	0.000	0.000
Nominal Interest Rate	0.000	0.002	0.002
Yield Spread	0.017	0.010	0.018

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

23

Table 3. Welfare Costs of Gap Fluctuations (1960-1999)
(percent average consumption)

		$\varphi = 1$	$\varphi = 5$	$\varphi = 10$
$\sigma = 1$				
	μ			
	0.10	0.082	0.289	0.559
	0.25	0.072	0.255	0.493
	0.40	0.064	0.228	0.440
$\sigma = 5$				
	μ			
	0.10	0.252	0.430	0.681
	0.25	0.240	0.410	0.633
	0.40	0.230	0.393	0.595
$\sigma = 10$				
	μ			
	0.10	0.477	0.641	0.871
	0.25	0.455	0.624	0.837
	0.40	0.437	0.610	0.811

Note: Calibration $a = 1$. The average output consumption ratio is 1.71. The data was filtered for the period 50:1-99:4 using a fifth order polynomial in time (e.g. Hall (1987).) Welfare computations cover the sample period 60:1-99:4.

Table 4 : Costs(−) and Benefits(+) of Boom/Recession Episodes
(percent of one year's consumption)

	72:2-78:1			78:2-84:1		
	Boom	*Reces.*	*Net*	*Boom*	*Reces.*	*Net*
Baseline ($\sigma = 1$, $\varphi = 5$, $\mu = 0.25$)	0.8	-5.7	-4.9	1.0	-4.4	-3.4
High Elasticity ($\sigma = 1$, $\varphi = 1$, $\mu = 0.25$)	0.9	-3.4	-2.5	0.8	-3.0	-2.2
Low Elasticity ($\sigma = 10$, $\varphi = 10$, $\mu = 0.25$)	0.2	-6.2	-6.0	0.3	-9.0	-8.7

Note: See Table 3.

25

Appendix

As shown in the text, the effect on welfare of a deviation from the frictionless level of employment can be approximated by

$$\Delta_t \simeq \left(\overline{U}_{c,t}Z_t + \overline{U}_{n,t}\right)\ (N_t - \overline{N}_t) + \frac{1}{2}\ \left(\overline{U}_{cc,t}\frac{\partial \overline{C}_t}{\partial \overline{Y}_t}Z_t + \overline{U}_{nn,t}\right)\ (N_t - \overline{N}_t)^2 + o(\|a\|^3)$$

where $o(\|a\|^n)$ represents terms that are of order higher than n^{th}, in the bound $\|a\|$ on the amplitude of the relevant shocks. A straightforward manipulation yields

$$\begin{aligned}
\Delta_t\ &\simeq\ \overline{U}_{c,t}\overline{Y}_t\ \left(1 + \frac{\overline{U}_{n,t}}{\overline{U}_{c,t}Z_t}\right)\left(\frac{N_t - \overline{N}_t}{\overline{N}_t}\right) + \\
&\quad \frac{1}{2}\left(\frac{\overline{U}_{cc,t}\overline{C}_t}{\overline{U}_{c,t}}\overline{\eta}_{cy,t} + \frac{\overline{U}_{n,t}}{\overline{U}_{c,t}Z_t}\frac{\overline{U}_{nn,t}\overline{N}_t}{\overline{U}_{n,t}}\right)\left(\frac{N_t - \overline{N}_t}{\overline{N}_t}\right)^2
\end{aligned}$$

We now make use of the second order approximation of relative deviations in terms of log deviations:

$$\frac{X - \overline{X}}{\overline{X}} = \widetilde{x} + \frac{1}{2}\ \widetilde{x}^2 + o(\|a\|^3)$$

where $\widetilde{x} \equiv \log\left(\frac{X}{\overline{X}}\right)$. Hence, the previous expression for Δ_t can be written as

$$\Delta_t \simeq \overline{U}_{c,t}\overline{Y}_t\left\{\left(1 - \frac{1}{1+\mu}\right)\left(\widetilde{n}_t + \frac{1}{2}\ \widetilde{n}_t^2\right) - \frac{1}{2}\ \left(\sigma\overline{\eta}_{cy,t} + \frac{\varphi}{1+\mu}\right)\ \widetilde{n}_t^2\right\},$$

where $\overline{\eta}_{cy,t} = \frac{\partial \overline{C}_t}{\partial \overline{Y}_t}\frac{\overline{Y}_t}{\overline{C}_t}$ is the elasticity of consumption with respect to output in the frictionless equilibirum, and we have used that $MPN_t = \frac{Y_t}{N_t} = Z_t$, $\frac{\overline{U}_{n,t}}{\overline{U}_{c,t}}/\overline{MPN}_t = GAP = \frac{1}{1+\mu}$, $\frac{\overline{U}_{cc,t}\overline{C}_t}{\overline{U}_{c,t}} = \sigma$, and $\frac{\overline{U}_{nn,t}\overline{N}_t}{\overline{U}_{n,t}} = \varphi$. Finally, from the previous expression it is straightforward to obtain the expression for Δ_t that is in the main text:

$$\Delta_t \approx \frac{\overline{U}_{c,t}\overline{Y}_t}{(1+\mu)}\ \left\{\mu\ \widetilde{n}_t\ -\frac{1}{2}\left[\varphi + \sigma(1+\mu)\overline{\eta}_{cy,t} - \mu\right]\ \widetilde{n}_t^2\right\}$$

26

REFERENCES

References

[1] Alexopolous, Michelle, (2001), "Unemployment and the Business Cycle," mimeo, University of Toronto.

[2] Altonji, Joseph G., (1986), "Intertemporal Substitution in Labor Supply: Evidence from Micro Data," *Journal of Political Economy,* 94, June, S176-S215.

[3] Ball, Lawrence and David Romer, (1989), "Are Prices Too Sticky?" *Quarterly Journal of Economics*, 1987, 104, 507-524.

[4] Barlevy, Gadi, 2000, "Evaluating Costs of Business Cycles of Business Cycles in Models of Endogenous Growth", mimeo.

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

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

[7] Basu, S. and M. Kimball (1997), "Cyclical Productivity with Unobserved Input Variation", NBER WP 5915.

[8] Baxter, M. and Robert King (1991), "Productive Externalities and Business Cycles", Discussion Paper no 53. Federal Reserve Bank of Minneapolis, Institute for Empirical Macroeconomics, November 1991.

[9] Beaudry, Paul and Carmen Pages (2001): "The Cost of Business Cycle and the Stabilization Value of Unemployment Insurance," *European Economic Review* 45, no. 8, 1545-1572.

[10] Bernanke, Ben and Ilyan Mihov (1998), "Measuring Monetary Policy," *Quarterly Journal of Economics, CXIII, 869-902.*

[11] Blanchard Olivier, J. and Nobuhiro Kiyotaki, (1987), "Monopolistic Competition and the Effects of Aggregate Demand," *American Economic Review,* LXXVII, 647-66.

[12] Burnside, Craig, and Martin Eichenbaum, (1996), "Factor Hoarding and the Business Cycel," *American Economic Review, 40, 861-69.*

27

[13] Card, David, (1994), "Intertemporal Labor Supply: An Assessment," in Frontiers of Econometrics, Christopher Sims, editor.

[14] Christiano, Lawrence, Martin Eichenbaum and Charle Evans, (1997), "Sticky Prices and Limited Participation Models: A Comparison," *European Economic Review,* 41, 1201-1249.

[15] Christiano, Lawrence, Martin Eichenbaum and Charle Evans, (1999), "Monetary Policy Shocks: What Have we Learned and to What End?," *Handbook of Macroeconomics, Vol.1A,.*Amsterdam North-Holland.

[16] Christiano, Lawrence, Martin Eichenbaum and Charle Evans, (2001), "Nominal Rigidities and the Effects of a Shock to Monetary Policy." mimeo Northwestern University.

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

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

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

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

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

[22] Gordon, Robert J. (1973), "The Welfare Cost of Higher Unemployment," *Brookings Papers on Economic Activity no. 1,* 133-195. .

[23] Hall, Robert, E., (1988), "Intertemporal Substitution in consumption," *Journal of Political Economy, 96 (2), 339-357.*

[24] Hall, Robert, E. (1997): "Macroeconomic Fluctuations and the Allocation of Time", *Journal of Labor Economics*, vol.15,no.1,pt.2. 1997, S223-S250.

[25] Holland, A. and A. Scott, (1998), "Determinants of U.K. Business Cycles," *Economic Journal,* 108, 1067-1092.

28

[26] King, Robert and Sergio Rebelo, (1999):"Resuscitating Real Business Cycles", *The Handbook of Macroeconomics, Vol 1B,* Amsterdam North-Holland.

[27] Lucas, Robert E., (1987), *Models of Business Cycles*

[28] MacCurdy, Thomas, (1981), "An Empirical Model of Labor Supply in a Life-Cycle Setting," *Journal of Political Economy, 89,* December, 1059-85.

[29] Mulligan, Casey B.(1998), "Substitution Over Time: Another Look at Life-Cycle Labor Supply," *NBER MacroAnnual*, vol.13, Ben Bernanke and Julio Rotemberg, editors.

[30] Mulligan, Casey B. (2002): "A Century of Labor-Leisure Distortions," NBER WP#8774.

[31] Pencavel, John, (1998), "The Market Work Behavior and Wages of Women, 1975-94," *Journal of Human Resources,* 33, 771-804.

[32] Pencavel, John, "A Cohort Analysis of the Association Between Work Hours and Wages Among Men, *Journal of Human Resources,* forthcoming.

[33] Rotemberg, Julio, and Michael Woodford (1999), "The Cyclical Behavior of Prices and Costs," *The Handbook of Macroeconomics, Vol 1B,* Amsterdam North-Holland.

[34] Sbordone, Argia, M. (1999), "Prices and Unit Labor Costs: A New Test of Price Stickiness". mimeo, Rutgers University.

[35] Sbordone, Argia, M., (2000), "An Optimizing Model of Wage and Price Dynamics," mimeo Rutgers University.

[36] Staiger, Douglas, James Stock and Mark Watson (2001), "Prices, Wages and the U.S. NAIRU in the 1990s," mimeo.

[37] Woodford, M. (1999): "Inflation Stabilization and Welfare" NBER Working Paper # 5684.
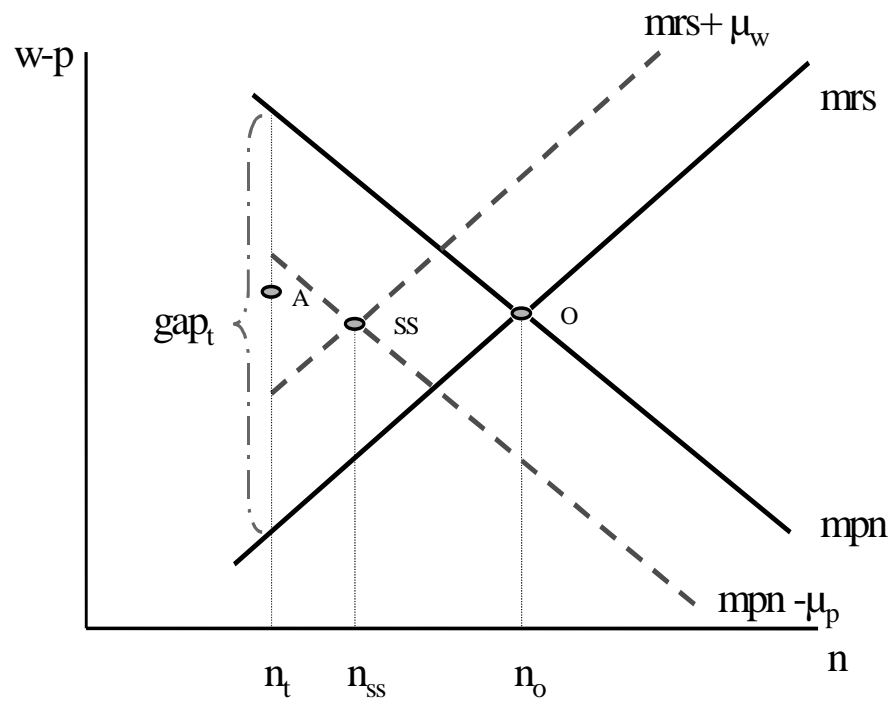
Figure 1. The Gap: A Diagrammatic Exposition

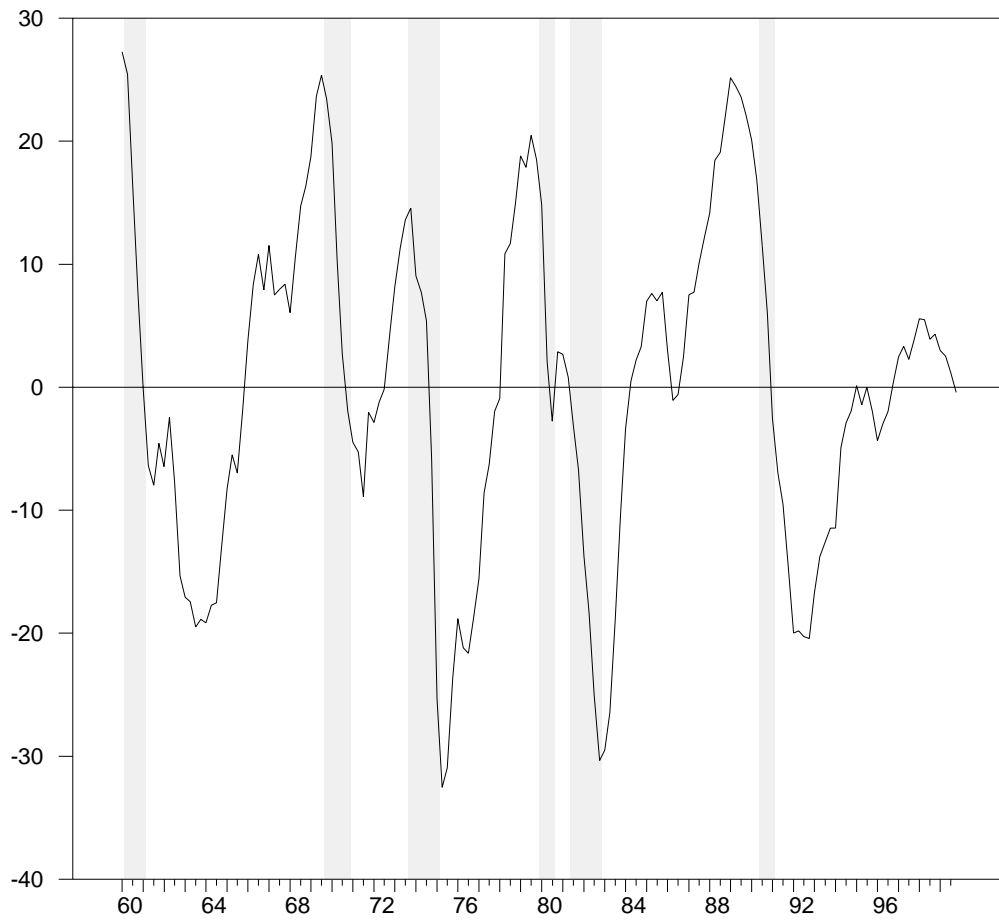# Figure 2. The Gap
Baseline Calibration (σ=1, φ=5)

Figure 3. The Gap and the Wage Markup
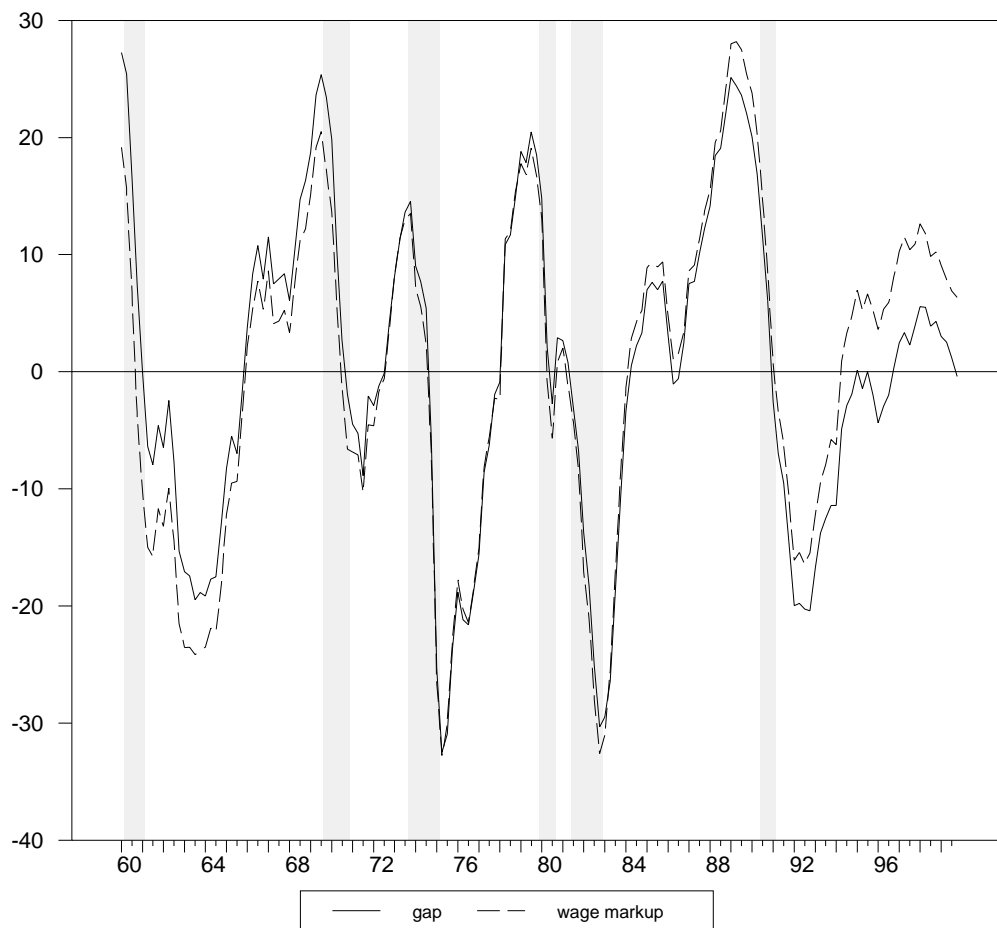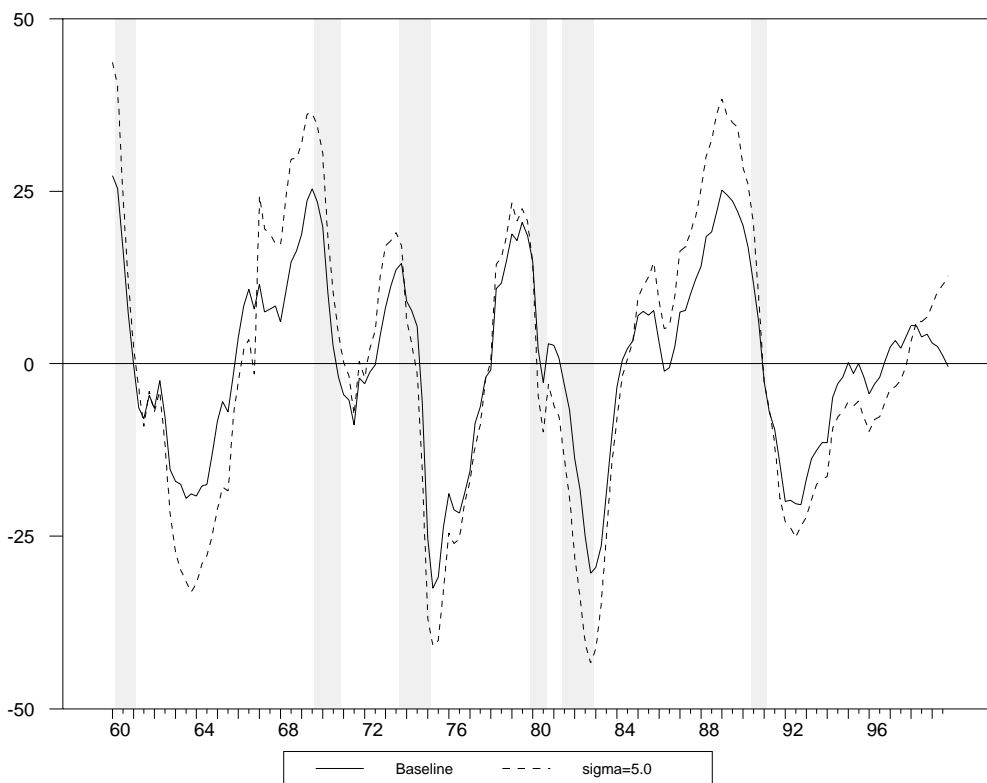## *Baseline Calibration*

Figure 4. The Gap under Alternative Calibrations

a. High Risk Aversion ($\sigma=5$)
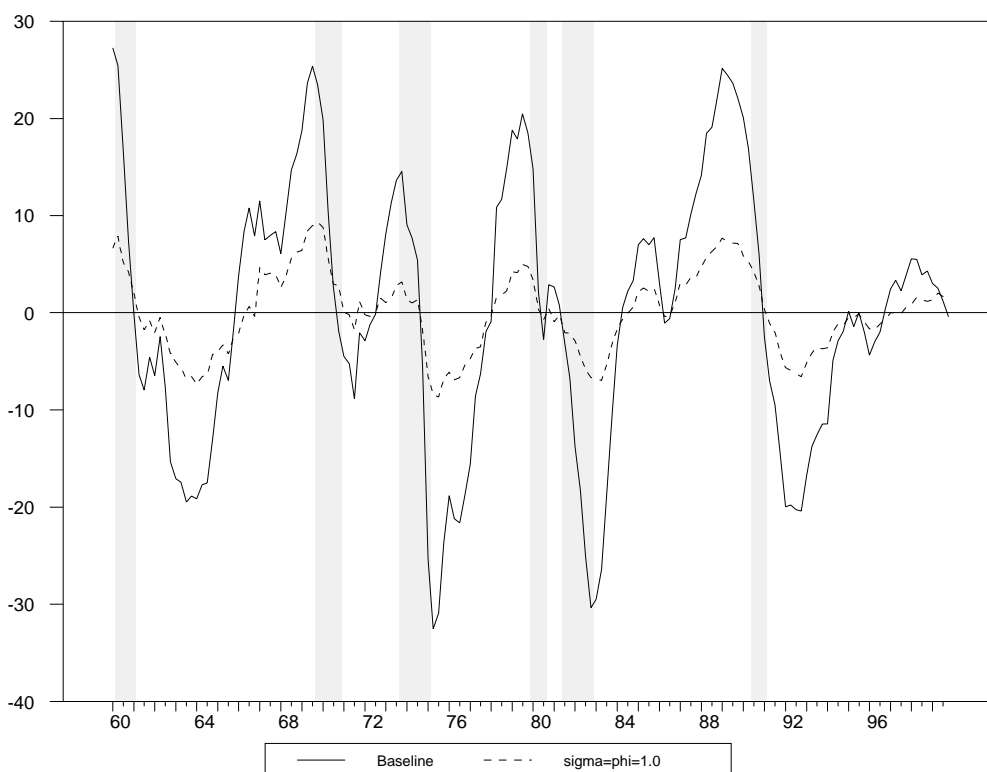

b. High Labor Supply Elasticity ($\varphi=1$)

Figure 5 . The Gap under Alternative Marginal Cost Measures
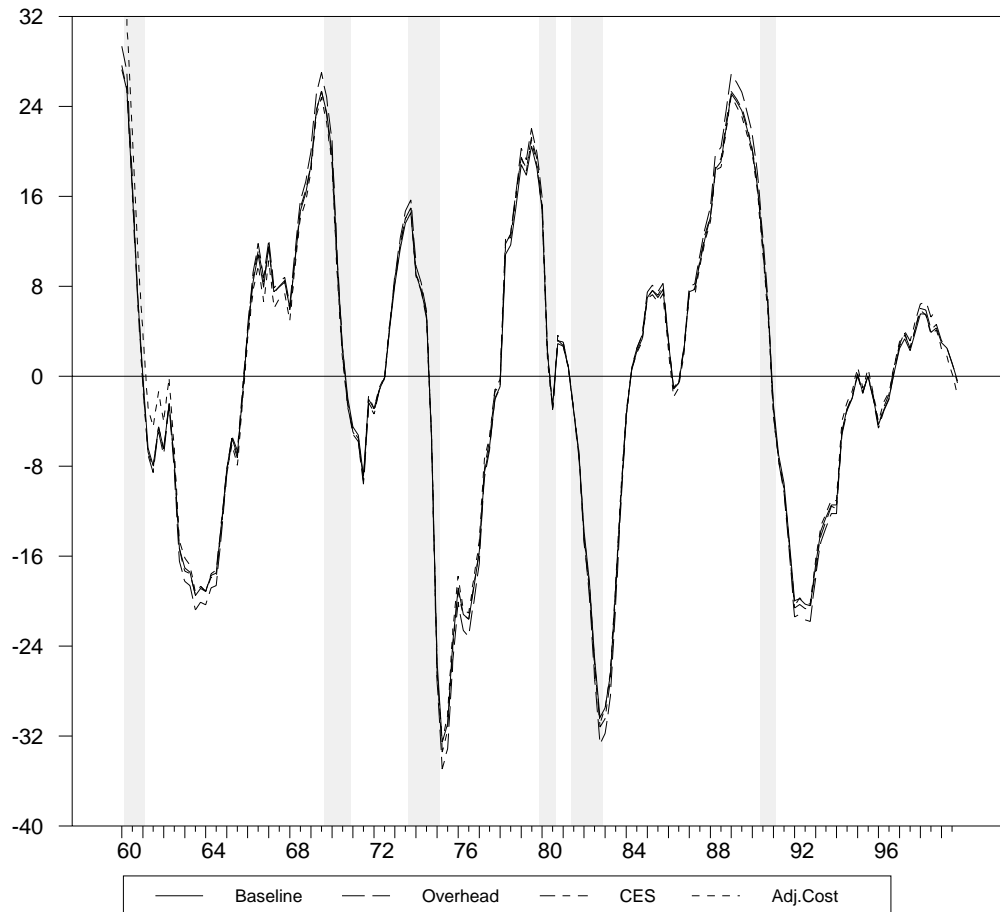## *Baseline Calibration*

Figure 6. The Gap under Non-Separable Preferences
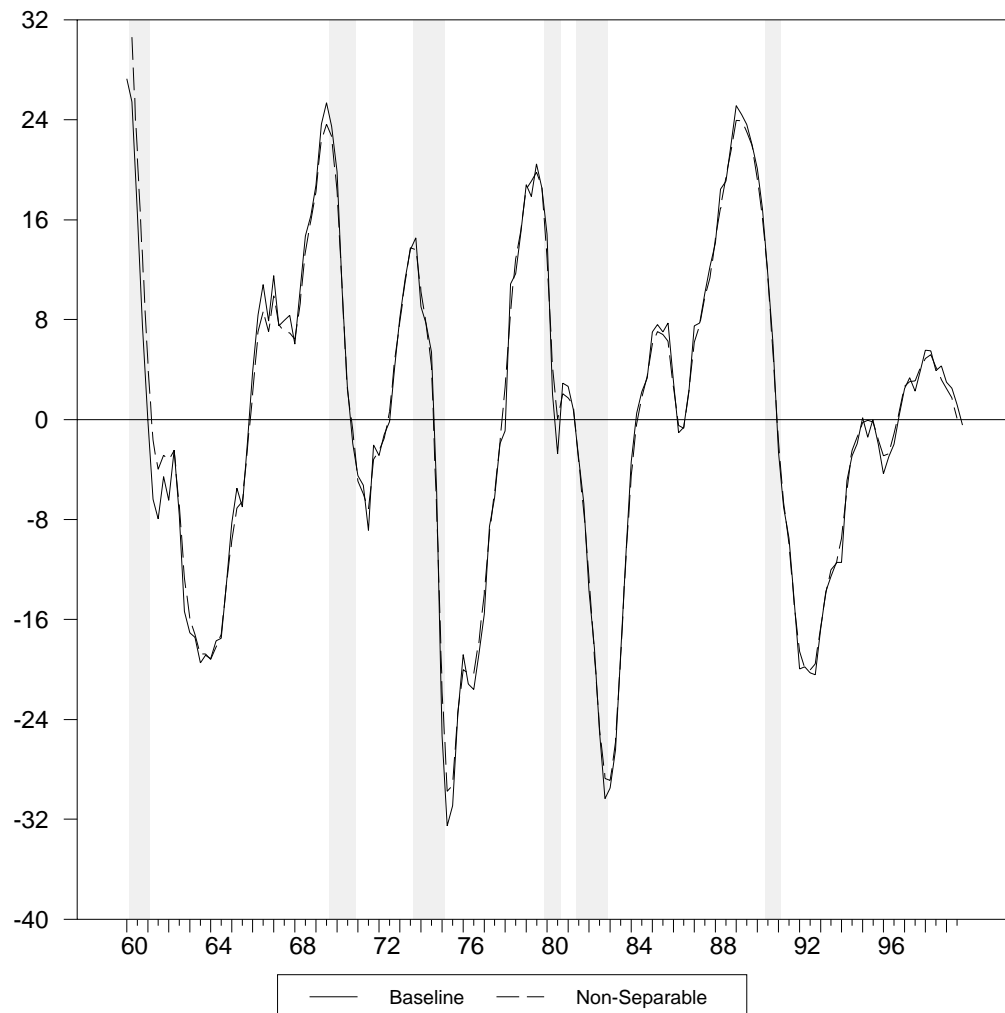## *Parameter Values: $\sigma=1$, $\varphi=5$, $\beta=0.99$, $b=0.8$*

Figure 7. Dynamic Effects of Monetary Policy Shocks
Baseline Calibration

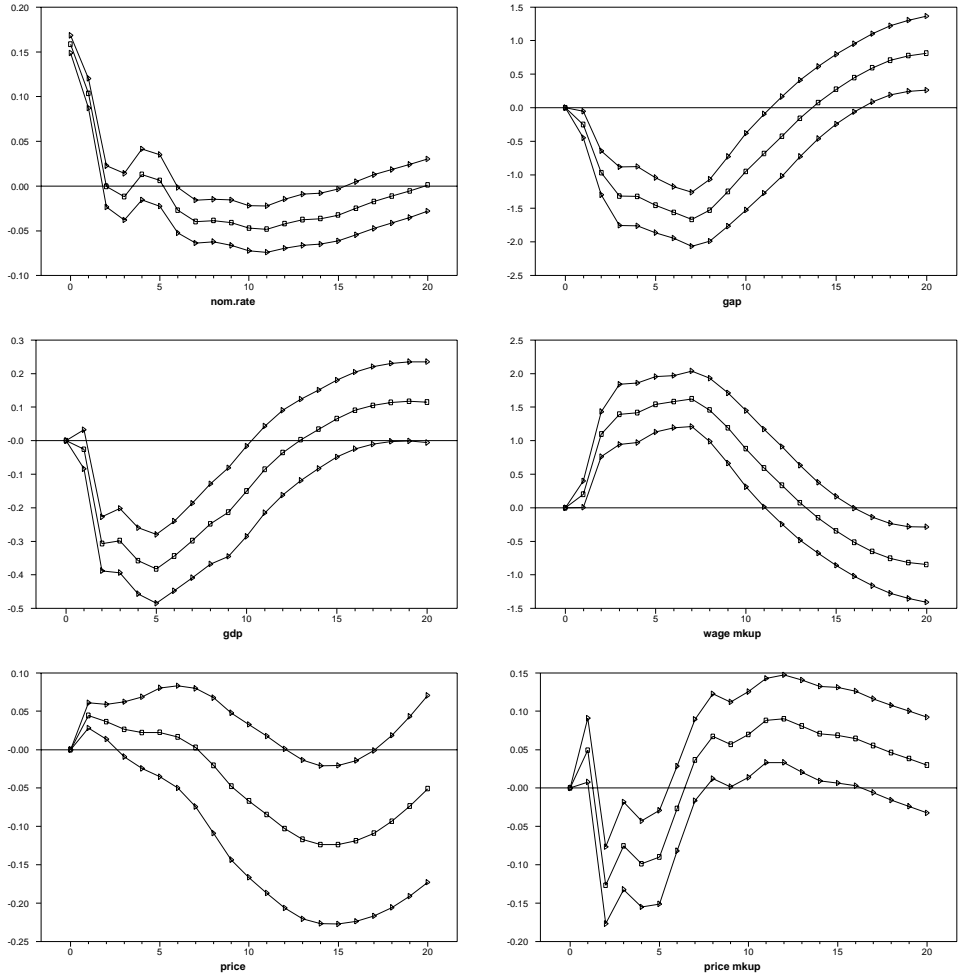# Figure 8. Theory-based vs. CBO based Output Gap
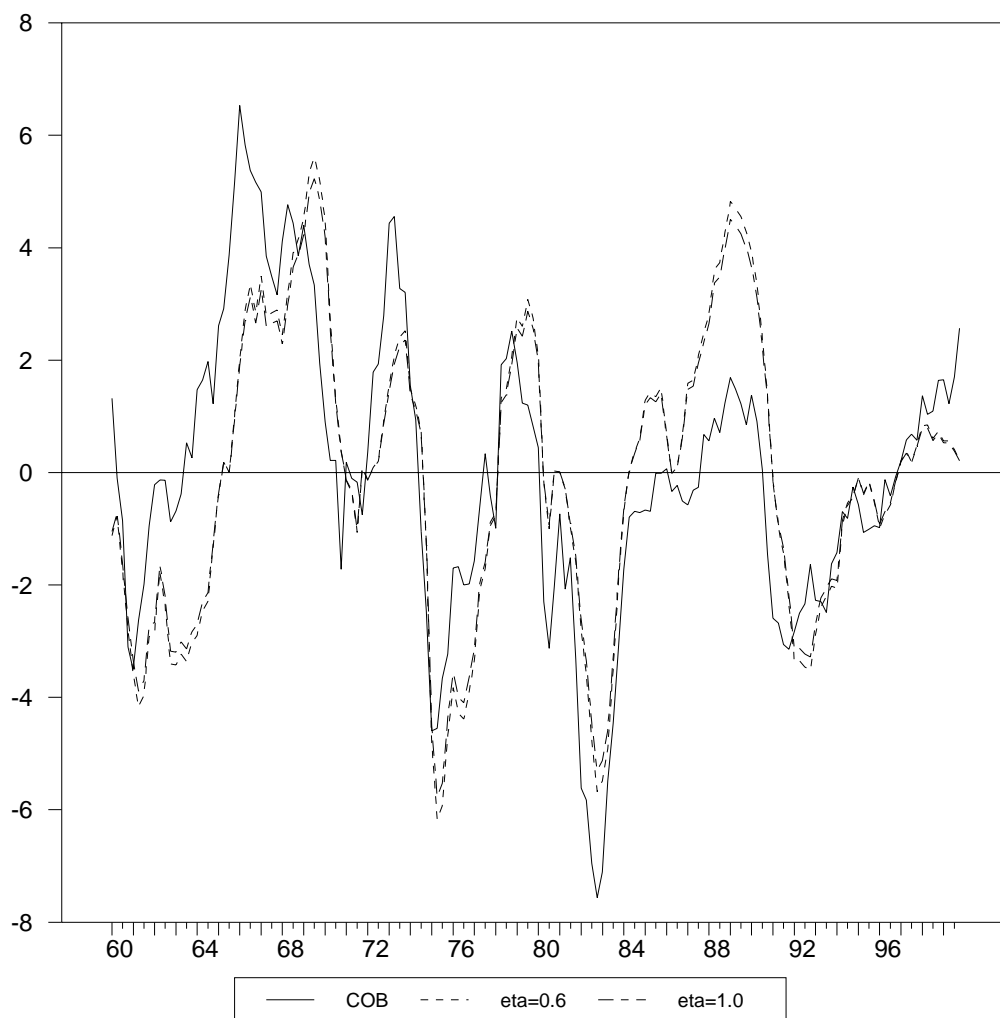### Baseline Calibration (σ=1, φ=5, μ=0.25)

Figure 9. The Welfare Effects of Postwar U.S. Fluctuations
Baseline Calibration ($\sigma=1, \varphi=5, \mu=0.25$)