IS THERE A TRADE-OFF BETWEEN INFLATION AND OUTPUT STABILIZATION?

ALEJANDRO JUSTINIANO, GIORGIO E. PRIMICERI, AND ANDREA TAMBALOTTI

Abstract. We find that the answer is no, in an estimated DSGE model of the U.S. economy in which exogenous movements in workers’ market power are not a major driver of observed economic fluctuations. If they are, the tension between the conflicting stabilization objectives of monetary policy increases, but with negligible effects on the equilibrium behavior of the economy under optimal policy.

1. Introduction

Between 1954 and 2009, average GDP growth in the U.S. was 3.2 percent per quarter at an annual rate, with a standard deviation of 3.8 percent. The origin of these fluctuations in economic activity is still a matter of debate. Even more controversial is the extent to which these fluctuations reflect movements in the economy’s efficient frontier, or inefficient departures from it. This distinction matters because stabilization policy can only play a useful role in the second case.

We make two contributions to this debate. First, we show that changes in the economy’s degree of inefficiency and movements of its efficient frontier were both important factors behind the observed fluctuations in U.S. output over the post-war period. Second, we compute the economy’s counterfactual evolution under optimal monetary policy and find that it involves a fairly weak trade-off between real and nominal stabilization. This result depends on the prior that exogenous movements in the competitiveness of the labor market are not a fundamental driver of macroeconomic fluctuations. If they are, the policy trade-off becomes steeper, but with only minor repercussions on the optimal equilibrium.

Date: First version: December 2010. This version: July 2012. This paper is a substantially revised and extended version of an earlier manuscript titled “Potential and Natural Output.” We thank Jose Dorich, Gauti Eggertsson, Chris Erceg, Stefano Eusepi, Andrea Ferrero, Jordi Gali, John Leahy, Pau Rabanal, Ricardo Reis, Juan Rubio-Ramirez, Frank Smets, Lars Svensson, Carl Walsh, John Williams, Mike Woodford and participants in several conferences and seminars for comments and suggestions, Ging Cee Ng for superb research assistance and especially Vasco Cúrdia for generously sharing his code for the computation of optimal equilibria. Giorgio Primiceri thanks the Alfred P. Sloan Foundation for financial support. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Chicago, New York or the Federal Reserve System.
Our results are based on an estimated dynamic stochastic general equilibrium (DSGE) model, which is built around neoclassical growth foundations, but assumes that firms and workers enjoy some monopoly power and cannot set prices freely every period. Relative to an environment with perfect competition (and flexible prices), these New-Keynesian features distort the economy’s equilibrium away from the efficient allocation. This distortion manifests itself in the form of markups of goods prices over nominal marginal costs and of real wages over the marginal rate of substitution between consumption and leisure. These markups vary over time due to exogenous movements in market competitiveness, and to the stickiness of prices and wages.

We summarize the evolution of these distortions by measuring the distance between actual and potential GDP, which we define as the level of GDP that would be observed if markups were constant at their steady state level. Since steady state markups are positive, potential GDP fluctuates around the same balanced growth path as actual GDP and hence it is inefficiently low. However, we show that its log-linear dynamics are very similar to those of efficient GDP, which would obtain under perfect competition and thus zero markups, and whose balanced growth path therefore lies above that of actual and potential GDP. Given the very similar dynamics of potential and efficient output, the output gap, which we define as the difference between actual and potential GDP, is a useful summary statistic of the movements of the economy away from its efficient frontier. We focus on the gap between actual and potential, rather than efficient, output because potential is a more natural reference point for monetary policy. The reason is that monetary policy is neutral in the long run, and thus it cannot affect the steady state distance between actual and efficient GDP. According to our estimates, the output gap is pro-cyclical and often quite large, with a standard deviation of 2.5 percentage points. In light of this evidence, we conclude that movements in the economy’s degree of inefficiency were an important factor behind fluctuations in the post-war U.S. economy.

Does the existence of inefficient fluctuations imply that stabilization policy was suboptimal over our sample? In general, the answer is no, since the model economy features multiple distortions that result in a complex trade-off between the stability of the output gap and other policy objectives. Most notably, the stabilization of the output gap produces dispersion in the cross-section of prices and wages, which in the aggregate manifests itself as inflation. For
example, the exogenous fluctuations in desired markups due to changes in market competitiveness give agents with the chance to adjust their prices a reason to do so, even when the output gap is zero, thus creating a discrepancy between newly set prices and the existing ones. Price dispersion, in turn, produces dispersion in markups and in the supply of goods and labor, which is inefficient because workers and firms are identical and the technology that aggregates their inputs is concave. As a result, producing the potential level of output in the actual economy requires more work than in the counterfactual economy with stable markups. Stated more formally, the full potential allocation, in which markups are constant across agents and time, is infeasible in our model. Therefore, policymakers must trade-off the stabilization of output around potential with that of price and wage inflation.\footnote{In environments with flexible wages, no markup shocks and no capital accumulation, the stabilization of the output gap and of aggregate markups are equivalent and produce no price dispersion, thus delivering the efficient allocation. This is what Blanchard and Gali (2007) call the “divine coincidence”. See Woodford (2003) or Gali (2008) for a textbook treatment. In our environment, output gap and markup stabilization are not equivalent, due to the presence of capital accumulation. As a result, the policy tradeoff is not only between output and inflation stabilization, but also involves the composition of demand between consumption and investment.}

To evaluate quantitatively the significance of this trade-off in our estimated model, we compute its optimal allocation, focusing in particular on optimal output. This is the counterfactual level of output that would have been observed in the post-war U.S. economy if the nominal interest rate had been set so as to maximize the utility of the model’s representative agent, rather than following the interest rate rule we estimate. The main finding of this exercise, and the central result of the paper, is that optimal and potential output move closely together, and that this stabilization of the output gap is roughly consistent with the stabilization of wage and price inflation. The surprising implication is that stabilization policy appears to face a negligible trade-off among its three main objectives and that much of the inefficient variation in output uncovered by our estimates could have been eliminated, increasing welfare at the same time.

The key factor behind the absence of a trade-off in our baseline specification is that the estimated contribution of wage markup shocks to macroeconomic fluctuations is rather small both at high and low frequencies. In our baseline, at high frequencies, this result is driven by our approach to wage measurement. Unlike in most empirical DSGE exercises, exemplified for instance by the seminal work of Smets and Wouters (2007), we match the model’s wage variable to two measures of hourly labor income, allowing for errors in their measurement, along the lines of Boivin and Giannoni (2006a) (see also Gali, Smets, and Wouters (2007)).
As a result, most of the high frequency variation that characterizes the individual series on compensation used in estimation is not interpreted as implausibly large variation in the monopoly power of workers from one quarter to the next. It is instead recognized as measurement error or, more generally, as a mismatch between the data and the model’s wage concept.

This shift in the interpretation of the factors behind the observed high frequency movements in wages—from markup shocks to measurement errors—produces a model that is economically more plausible, empirically more successful, and with radically different normative implications. Indeed, in models estimated with only one wage series, the extreme volatility of wage markup shocks implies that optimal output is significantly more volatile than actual output. This destabilization of output is the price policymakers must pay to obtain the desired stabilization of wages, whose dispersion is extremely costly for the representative agent.

We estimate a model of this kind in section 6.1. In contrast, this extreme tension between output and wage stabilization is virtually absent in the baseline model with multiple wage indicators, where the variation in wage markups is small. In fact, the deviation of the optimal from the potential allocation is altogether minimal in this model, leading to our conclusion that the conflict among policymakers’ objectives is negligible.

At low frequencies, the contribution of wage markup shocks to fluctuations is small because our prior assumes that most of the exogenous variation in labor supply identified by the model is due to shifts in households’ attitudes towards work in the market. If these secular changes in labor supply were interpreted instead as originating exclusively from a trend in workers’ monopoly power—an hypothesis that our empirical procedure cannot rule out—inferring on the movements of the economy’s efficient frontier, and therefore of the output gap, would change significantly (Sala, Söderström, and Trigari (2010)).

Even under this extreme alternative scenario, however, the properties of the equilibrium under optimal policy, and of optimal output in particular, change little with respect to the baseline specification. This is because wage dispersion is very costly, and therefore a stable path for wage inflation is always optimal. As a result, quantities—hours and output in particular—bear the brunt of the adjustment to labor supply shocks in the optimal equilibrium, regardless of these shocks’ origin.

---

2 The contribution of these changes at business cycle frequencies is small, regardless of their interpretation.
This finding is worth highlighting, because it reminds us that the common distinction between “efficient” and “inefficient” shocks—shocks that do or do not affect a model’s efficient equilibrium—can be a misleading guide to optimal policy. More specifically, it is not true that policy should accommodate the shocks that shift the efficient frontier, such as those to tastes and technology, but offset the ones that leave that frontier unchanged, such as those to market power (Chari, Kehoe, and McGrattan (2009)). In our model, in fact, the economy’s response to labor supply shocks under optimal policy does not change much whether those shocks come from changes in workers’ tastes or in their monopoly power, even if the response of the efficient economy does.

This paper is related to a large literature on the estimation of DSGE models (e.g. Rotemberg and Woodford (1997), Lubik and Schorfheide (2004), Boivin and Giannoni (2006b), Smets and Wouters (2007)). Some of these papers explicitly assume that monetary policy reacts to the difference between actual and potential output, but do not focus on the model’s predictions regarding the behavior of the output gap. A number of studies, such as Levin, Onatski, Williams, and Williams (2005), Nelson (2005), Andrés, López-Salido, and Nelson (2005) and Edge, Kiley, and Laforte (2008), have tackled this issue more directly, but have found it difficult to obtain model-based estimates of the output gap with reasonable cyclical properties. In comparison, our DSGE output gap is more consistent with conventional views of the business cycle, and thus closer to the estimates of Sala, Soderstrom, and Trigari (2008), who were the first to obtain a cyclical output gap in an estimated DSGE model, and Gali, Smets, and Wouters (2011). In this positive dimension, our work is closely related to Gali, Gertler, and Lopez-Salido (2007), who also measure the extent of inefficient fluctuations, but do so through the lens of the labor wedge, rather than of the output gap.

Although the cyclical properties of the output gap are interesting, our main contribution is instead to interpret this object from a more explicit normative perspective. In this dimension, we make contact with the optimal policy literature in medium-scale DSGE models (e.g. Levin, Onatski, Williams, and Williams (2005), Schmitt-Grohe and Uribe (2004) and Schmitt-Grohé and Uribe (2007)). As these authors, we find that nominal dispersion is key for the normative implications of the model. Unlike them, however, we find virtually no tension between inflation and output gap stabilization, once we recognize that wage markup shocks are likely to be small. The important role of wage markup shocks and labor supply shocks more in general relates our work to Chari, Kehoe, and McGrattan (2009) and Sala,
Söderström, and Trigari (2010), although they are not concerned with the characterization of optimal policy.

The rest of the paper is organized as follows. Section 2 provides the details of the theoretical model and section 3 describes the approach to measurement and inference. Sections 4 and 5 present our estimates of potential and optimal output, while section 6.1 analyzes the role of labor supply shocks in the results. Section 7 concludes.

2. THE MODEL ECONOMY

This section outlines our baseline model of the U.S. business cycle, which is similar to Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007). It is a medium-scale DSGE model with a neoclassical growth core, augmented with several shocks and “frictions”—departures from the simplest assumptions on tastes, technology and market structure—now common in the literature.

The economy is populated by five classes of agents: producers of a final good, intermediate goods producers, households, employment agencies and a government. We now present their optimization problems.

2.1. Final good producers. At every point in time \( t \), perfectly competitive firms produce the final good \( Y_t \) by combining a continuum of intermediate goods \( \{Y_t(i)\}_{i \in [0,1]} \), according to the technology

\[
Y_t = \left[ \int_0^1 Y_t(i) \frac{i}{1 + \Lambda_{p,t}} \, di \right]^{1 + \Lambda_{p,t}}.
\]

Profit maximization and the zero profit condition imply that the price of the final good, \( P_t \), is a CES aggregate of the prices of the intermediate goods, \( \{P_t(i)\}_i \)

\[
P_t = \left[ \int_0^1 P_t(i) \frac{1}{\Lambda_{p,t}} \, di \right]^{-\frac{1}{\Lambda_{p,t}}},
\]

and that the demand function for intermediate good \( i \) is

\[
Y_t(i) = \left( \frac{P_t(i)}{P_t} \right)^{-\frac{1 + \Lambda_{p,t}}{\Lambda_{p,t}}} Y_t.
\]

The curvature of the aggregator \( \Lambda_{p,t} \) determines the degree of substitutability across intermediate goods in the production of the final good and hence the elasticity of demand for
each of these intermediates, as shown in (2.1). We model \( \Lambda_{p,t} \) as an exogenous stochastic process

\[
\log (1 + \Lambda_{p,t}) \equiv \lambda_{p,t} = (1 - \rho_p) \lambda_p + \rho_p \lambda_{p,t-1} + \varepsilon_{p,t},
\]

driven by innovations \( \varepsilon_{p,t} \) distributed i.i.d.\( N(0, \sigma_p^2) \). We refer to these innovations as price markup shocks, since, according to Lerner’s formula, \( \Lambda_{p,t} \) is the desired net markup of price over marginal cost for intermediate firms. Therefore, \( \Lambda_{p,t} \) is also a measure of the (lack of) competitiveness in the intermediate goods market and its exogenous movements are one of the forces driving the economy away from its efficient frontier.

2.2. Intermediate goods producers. A monopolist produces the intermediate good \( i \) using the production function

\[
Y_t(i) = A_t^{1-\alpha} K_t(i)^\alpha L_t(i)^{1-\alpha} - A_t F,
\]

where \( K_t(i) \) and \( L_t(i) \) denote the amounts of capital and labor employed by firm \( i \). Both of these inputs are homogenous and \( F \) is a fixed cost of production, chosen so that profits are zero in steady state (see Rotemberg and Woodford, 1995 or Christiano, Eichenbaum, and Evans, 2005).

\( A_t \) represents exogenous labor-augmenting technological progress or, equivalently, a neutral technology factor. The level of neutral technology is non-stationary and its growth rate \( (z_t \equiv \Delta \log A_t) \) follows an AR(1) process

\[
z_t = (1 - \rho_z) \gamma + \rho_z z_{t-1} + \varepsilon_{z,t},
\]

with \( \varepsilon_{z,t} \) i.i.d.\( N(0, \sigma_z^2) \).

As in Calvo (1983), every period a fraction \( \xi_p \) of intermediate firms cannot optimally choose their price, but reset it according to the indexation rule

\[
P_t(i) = P_{t-1}(i) \pi_t^{\beta p} \pi_t^{1-\epsilon_p},
\]

where \( \pi_t \equiv \frac{P_t}{P_{t-1}} \) is gross inflation and \( \pi \) is its steady state. This indexation scheme implies no price dispersion in steady state. Therefore, the level of \( \pi \) is inconsequential for welfare, which allows us to abstract from the challenging question of the optimal level of inflation (see Schmitt-Grohe and Uribe (2010) for a recent survey.) In addition, full indexation produces a vertical Phillips curve in the long-run, so that the steady state level of output is independent
from that of inflation, regardless of policy. As a consequence, monetary policy cannot bring the economy closer to its efficient frontier on average, even if this shift might be desirable.

The remaining fraction of firms choose their price, $\tilde{P}_t(i)$, by maximizing the present discounted value of future profits

$$E_t \sum_{s=0}^{\infty} \xi_p \beta^s \frac{\Lambda_{t+s}}{\Lambda_t} \left\{ \left[ \tilde{P}_t(i) \left( \Pi_{j=0}^{s} \sigma_{t}^{j+1} \pi^{1-i_p} \right) \right] Y_{t+s}(i) - \left[ W_t L_t(i) + r^k_t K_t(i) \right] \right\},$$

subject to the demand function 2.1 and the production function 2.2. In this objective, $\Lambda_{t+s}$ is the marginal utility of nominal income of the representative household that owns the firm, while $W_t$ and $r^k_t$ are the nominal wage and the rental rate of capital.

2.3. Employment agencies. Firms are owned by a continuum of households, indexed by $j \in [0, 1]$. Each household is a monopolistic supplier of specialized labor, $L_t(j)$, as in Erceg, Henderson, and Levin (2000). A large number of competitive “employment agencies” combine these specialized types of labor into a homogenous labor input sold to intermediate firms, according to

$$L_t = \left[ \int_0^{1} L_t(j) \frac{1}{1+\lambda_{w,t}} dj \right]^{1+\frac{1}{\lambda_{w,t}}}. \tag{2.3}$$

As in the case of the final good production, the elasticity of this aggregator $\lambda_{w,t}$ corresponds to the desired markup of wages over households’ marginal rate of substitution between consumption and leisure. Unlike for the price markup shock, which is an AR(1) process, we assume instead that $\log \left( 1 + \lambda_{w,t} \right) \equiv \lambda_{w,t}$ is $i.i.d. N(0, \sigma^2_{w})$, for reasons explained in the next subsection. We refer to $\lambda_{w,t}$ as the wage markup shock.

Profit maximization by the perfectly competitive employment agencies implies the labor demand function

$$L_t(j) = \left( \frac{W_t(j)}{W_t} \right)^{-\frac{1+\lambda_{w,t}}{\lambda_{w,t}}} L_t, \tag{2.4}$$

where $W_t(j)$ is the wage paid by the employment agencies to the supplier of labor of type $j$, while

$$W_t = \left[ \int_0^{1} W_t(j) \frac{1}{\lambda_{w,t}} dj \right]^{-\lambda_{w,t}}$$

is the wage paid by intermediate firms for the homogenous labor input sold to them by the agencies.
2.4. **Households.** Each household $j$ maximizes the utility function

$$E_t \left\{ \sum_{s=0}^{\infty} \beta^s b_{t+s} \left[ \log (C_{t+s} - hC_{t+s-1}) - \varphi_t \frac{L_{t+s}(j)^{1+\nu}}{1+\nu} \right] \right\},$$

where $C_t$ is consumption and $h$ is the degree of habit formation. The disturbance to the discount factor $b_t$ is an *intertemporal preference shock* and follows the stochastic process

$$\log b_t = \rho_b \log b_{t-1} + \varepsilon_{b,t},$$

with $\varepsilon_{b,t} \sim i.i.d. N(0, \sigma_b^2)$. The disturbance to the disutility of labor $\varphi_t$ is instead an *intradimensional preference* or *labor supply shock*, as in Hall (1997). This shock enters households’ first order conditions for the optimal supply of labor in exactly the same way as the wage markup shock. As a consequence, these two disturbances are not separately identified in this model, when only using data on wages and total hours.\(^3\) However, the implications of these two shocks for the evolution of potential output differ markedly, as also pointed out by Chari, Kehoe, and McGrattan (2009) and Sala, Söderström, and Trigari (2010).

To disentangle these two disturbances, therefore, we model the labor supply shock as an AR(1) process

$$\log \varphi_t = (1 - \rho_\varphi) \varphi + \rho_\varphi \log \varphi_{t-1} + \varepsilon_{\varphi,t},$$

with $\varepsilon_{\varphi,t} \sim i.i.d. N(0, \sigma_\varphi^2)$. The autocorrelation of the taste shock $\varphi_t$, which distinguishes it from the *i.i.d.* wage markup shock $\Lambda_{w,t}$, reflects the prior view that taste shocks are a more plausible device to capture the effects on hours worked of changes in labor force participation and other low-frequency movements largely unrelated to the business cycle, but are nonetheless evident in our data. We will return to the observational equivalence of these two shocks and to its normative implications in sections 6.2, where we show that reinterpreting the estimated labor disutility shocks as markup shocks has a limited impact on the properties of optimal output.

Since technological progress is non stationary, utility is logarithmic to ensure the existence of a balanced growth path. Moreover, consumption is not indexed by $j$ because the existence of state contingent securities ensures that in equilibrium consumption and asset holdings are the same for all households.

\(^3\)Gali (2011) and Gali, Smets, and Wouters (2011) propose a reinterpretation of this same model with an explicit treatment of unemployment, in which the two shocks can be separately identified. For a DSGE model with similar implications, but alternative microfoundations of unemployment, see Christiano, Trabandt, and Walentin (2010).
As a result, the household’s flow budget constraint is

\[ P_t C_t + P_t I_t + T_t + B_t \leq R_{t-1} B_{t-1} + Q_t(j) + \Pi_t + W_t(j) L_t(j) + r^k_t u_t \bar{K}_{t-1} - P_t a(u_t) \bar{K}_{t-1}, \]

where \( I_t \) is investment, \( B_t \) is holdings of government bonds, \( R_t \) is the gross nominal interest rate, \( Q_t(j) \) is the net cash flow from household’s \( j \) portfolio of state contingent securities, \( \Pi_t \) is the per-capita profit accruing to households from ownership of the firms and \( T_t \) is lump-sum taxes and transfers.

Households own capital and choose the capital utilization rate, \( u_t \), which transforms physical capital into effective capital according to

\[ K_t = u_t \bar{K}_{t-1}. \]

Effective capital is then rented to firms at the rate \( r^k_t \). The cost of capital utilization is \( a(u_t) \) per unit of physical capital. We parameterize it as \( a(u_t) = \rho \frac{u^1+\chi^{-1}}{1+\chi} \), as in Levin, Onatski, Williams, and Williams (2005), such that in steady state, \( u = 1, a(1) = 0 \) and \( \chi \equiv \frac{a''(1)}{a'(1)} \). In the log-linear approximation of the model solution this curvature is the only parameter that matters for the dynamics.

The physical capital accumulation equation is

\[ \dot{K}_t = (1-\delta) \bar{K}_{t-1} + \mu_t \left( 1 - S \left( \frac{I_t}{I_{t-1}} \right) \right) I_t, \]

where \( \delta \) is the depreciation rate. The function \( S \) captures the presence of adjustment costs in investment, as in Christiano, Eichenbaum, and Evans (2005), and we parameterize it as \( S (I_t/I_{t-1}) = \frac{\zeta}{2} (I_t/I_{t-1} - e^\gamma)^2 \). In steady state, \( S = S' = 0 \) and \( S'' = \zeta > 0 \). This coefficient is also the only one that matters for the log-linear dynamics. The investment shock \( \mu_t \) is a source of exogenous variation in the efficiency with which the final good can be transformed into physical capital, and thus into tomorrow’s capital input. Justiniano, Primiceri, and Tambalotti (2011) show that this variation might stem from technological factors specific to the production of investment goods, as in Greenwood, Hercowitz, and Krusell (1997), but also from disturbances to the process by which these investment goods are turned into productive capital. The investment shock follows the stochastic process

\[ \log \mu_t = \rho_\mu \log \mu_{t-1} + \varepsilon_{\mu,t}, \]

where \( \varepsilon_{\mu,t} \) is \( i.i.d. N(0, \sigma^2_\mu) \).
As in Erceg, Henderson, and Levin (2000), every period a fraction \( \xi_w \) of households cannot freely set its wage, but follows the indexation rule

\[
W_t(j) = W_{t-1}(j) \left( \pi_{t-1} e^{\xi_{t-1}} \right)^{\xi_w} \left( \pi e^{\gamma} \right)^{1-\xi_w},
\]

which depends on the growth rate of neutral technology to guarantee the existence of a balance growth path. The remaining fraction of households chooses instead an optimal wage by maximizing their utility, subject to the labor demand function 2.4.

2.5. **Monetary and government policies.** When estimating the model and characterizing its positive properties, such as the behavior of potential output, we assume that the short-term nominal interest rate follows a feedback rule, of the type that has been found to provide a good description of actual monetary policy in the United States at least since Taylor (1993). Our specification of this policy rule features interest rate smoothing, a systematic response to deviations of annual inflation from a time varying inflation target, and to deviations of observed annual GDP growth \( (X_t/X_{t-4}) \) from its steady state level

\[
R_t = \left( \frac{R_{t-1}}{R} \right)^{\rho_R} \left[ \prod_{s=0}^{3} \frac{\pi_{t-s}}{\pi_t^{1/4}} \right]^{1/4} \left[ \frac{(X_t/X_{t-4})^{1/4}}{e^{\gamma}} \right]^{1-\rho_R} e^{\xi_{R,t}},
\]

where \( R \) is the steady state gross nominal interest rate and \( \xi_{R,t} \) is an *i.i.d.* \( N(0, \sigma^2_R) \) monetary policy shock.

The inflation target \( \pi^*_t \) evolves exogenously according to the process

\[
\log \pi^*_t = (1 - \rho_\pi) \log \pi + \rho_\pi \log \pi^*_{t-1} + \varepsilon_{\pi,t},
\]

with \( \varepsilon_{\pi,t} \sim i.i.d. N(0, \sigma^2_\pi) \). The primary role of this inflation target is to account for the very low frequency behavior of inflation (Ireland (2007)). In reduced form, these secular movements might reflect the slow evolution of policymakers’ beliefs and the consequent changes in the conduct of monetary policy, as suggested for instance by Cogley and Sargent (2005) and Primiceri (2006).

When characterizing optimal policy, this rule is ignored, and we assume that the central bank sets the interest rate so as to directly maximize the utility of the representative agent.
Fiscal policy is Ricardian. The government finances its budget deficit by issuing short term bonds. Public spending is determined exogenously as a time-varying fraction of output

$$G_t = \left(1 - \frac{1}{g_t}\right)Y_t,$$

where the government spending shock $g_t$ follows the stochastic process

$$\log g_t = (1 - \rho_g) \log g + \rho_g \log g_{t-1} + \varepsilon_{g,t},$$

with $\varepsilon_{g,t} \sim i.i.d. N(0, \sigma_g^2)$.

### 3. Model Solution and Estimation

This section briefly describes the solution and estimation of the model (see An and Schorfheide (2007) for a survey). First, we collect the first order conditions and constraints of the agents’ optimization problems in a system of rational expectation difference equations, which characterize the equilibrium. The real variables of this economy are non-stationary because of technological progress. Therefore, we rewrite the equilibrium conditions in terms of deviations of the real variables from the non-stationary technology process $A_t$. Let

$$(3.1) E_t [f (\zeta_{t+1}, \zeta_t, \zeta_{t-1}, \varepsilon_t, \theta)] = 0,$$

denote the collection of these re-normalized equilibrium conditions, in which $\zeta_t$, $\varepsilon_t$ and $\theta$ are the vectors of endogenous variables, exogenous i.i.d. disturbances and unknown structural coefficients respectively. To obtain an estimate of our two main objects of interest, potential and optimal GDP, (3.1) must also include the equilibrium conditions of the corresponding counterfactual economies with stable markups and optimal monetary policy, and $\zeta_t$ must also contain all the variables necessary to characterize the dynamics of these counterfactual economies, including potential and optimal output.

We then log-linearize (3.1) around the non-stochastic steady state and solve the resulting linear system of rational expectation equations by standard methods (for example Sims (2001)). This procedure yields the following system of transition equations

$$(3.2) \hat{\zeta}_t = G(\theta) \hat{\zeta}_{t-1} + M(\theta) \varepsilon_t,$$

where the “hat” denotes log deviations from the steady state, $\hat{\zeta}_t$ is an extended version of $\hat{\zeta}_t$ that also includes the expectational variables that are necessary to characterize the solution.
of the model, and $G(\theta)$ and $M(\theta)$ are conformable matrices whose elements are functions of $\theta$.

3.1. Data and Measurement. We estimate the model using eight series of U.S. quarterly data: the inflation rate, the nominal interest rate, the logarithm of per-capita hours, the log-difference of real per-capita GDP, consumption and investment, and \textit{two} measures of nominal hourly wage inflation.

The inflation rate is the quarterly log difference of the GDP deflator, while the nominal interest rate is the effective Federal Funds rate. We measure per-capita hours as the number of hours worked in the total economy, divided by the civilian non-institutional population (16 years and older).\textsuperscript{4} The series of hours for the total economy exhibits a less pronounced low frequency behavior than that for the non-farm business sector, because it accounts better for sectoral shifts, as shown by Francis and Ramey (2009). Real per-capita GDP is nominal GDP divided by population and the GDP deflator. The real series for per-capita consumption and investment are obtained in the same manner. Consumption corresponds to the sum of non-durables and services, while investment is constructed by adding consumer durables to total private investment, all in nominal terms.

As pointed out in the introduction, we match the wage inflation variable in the model, $\Delta \log W_t$, with two data series, following the methodology proposed by Boivin and Giannoni (2006a) and recently also adopted by Gali, Smets, and Wouters (2011). The first series is nominal compensation per hour in the total economy, from NIPA. The second measure is the “average hourly earnings of production and non-supervisory employees,” which is computed by the Bureau of Labor Statistics from the Establishment Survey, and is the one preferred by Galí (2011). We assume that both series represent an imperfect match to the concept of “wage” in the model, and capture this mismatch through a simple i.i.d. observation error, as follows

$$
\begin{bmatrix}
\Delta \log NHC_t \\
\Delta \log HE_t
\end{bmatrix}
= \begin{bmatrix}
1 \\
\Gamma
\end{bmatrix}
\Delta \log W_t
+ \begin{bmatrix}
e_{1,t} \\
e_{2,t}
\end{bmatrix},
$$

where $\Delta \log NHC_t$ and $\Delta \log HE_t$ denote the growth rate of the two measure of wages in the data, $\Gamma$ is a loading coefficients relating the second series to wage inflation in the model (the other loading is normalized to 1, as standard in factor analysis), and $e_{1,t}$ and $e_{2,t}$ are

\textsuperscript{4}We are grateful to Shawn Sprague, of the Bureau of Labor Statistics, for providing the data on hours and the labor share in the total economy. Breaks in the civilian population series due to census-based population adjustments are smoothed by splicing them uniformly over a 10-year window.
i.i.d. observation errors with distribution $N(0, \sigma_{e1}^2)$ and $N(0, \sigma_{e2}^2)$. We will return to the reasons for this approach to the measurement of wages, and to its positive and normative implications in section 6.1.

The estimation sample starts in 1964:II, due to limited availability of the wage data, and ends in 2009:IV. We do not demean or detrend any series.

3.2. **Bayesian Inference and Priors.** We characterize the posterior distribution of the model’s coefficients by combining the likelihood function with prior information. The likelihood function can be evaluated by applying the Kalman Filter. Conditional on the sample information, the Kalman filter and smoother can also be used to estimate the historical path of the model’s endogenous variables, $\{\tilde{\xi}_t\}_{t=1}^T$, which include potential and optimal output. In the rest of this section we briefly discuss the specification of the priors, which is reported in Table 1.

Two parameters are fixed using level information not contained in our dataset: the quarterly depreciation rate of capital ($\delta$) to 0.025 and the steady state ratio of government spending to GDP ($1 - 1/g$) to 0.2, which corresponds to the average value of $G_t/X_t$ in our sample. Also due to lack of identification, we set the steady state net wage markup to 25 percent. The priors on the other coefficients are fairly diffuse and broadly in line with those adopted in previous studies, such as Smets and Wouters (2007) and Del Negro, Schorfheide, Smets, and Wouters (2007).

The prior distribution of all but two persistence parameters is a Beta, with mean 0.6 and standard deviation 0.2. The two exceptions are the autocorrelation of TFP shocks—whose prior is centered at 0.4, since this process already includes a unit root—and the autocorrelation of the inflation target shock, which we fix at 0.995. Our view is that the exogenous movements of the inflation target should account for the very low frequency behavior of inflation. In reduced form, these secular movements might reflect the slow evolution of policymakers’ beliefs and the consequent changes in the conduct of monetary policy, as suggested for instance by Cogley and Sargent (2005) and Primiceri (2006).

The intertemporal preference, price and wage markup shocks are normalized to enter with a unit coefficient in the consumption, price inflation and wage equations respectively (see appendix A for details). The priors on the standard deviations of the innovations to these normalized shocks are quite disperse and chosen to generate volatilities for the variables they
impact directly broadly in line with those in the data. The covariance matrix of the vector of shocks is diagonal.

3.3. Posterior Estimates of the Parameters. Table 1 summarizes the posterior estimates of the parameters in our baseline specification. The data are quite informative about these parameters and the estimates we obtain are generally in line with those of previous studies. For this reason, and given the focus of our paper on the implications of these estimates for the inefficiency of the economy, and for optimal policy, we only briefly comment on the coefficients related to nominal rigidities and to monetary policy. The posterior distributions of the parameters $\xi_p$ and $\xi_w$ imply that prices and wages are re-optimized approximately every year and every three quarters respectively, while $t_p$ and $t_w$ indicate very low levels of backward indexation. Steady state inflation is about 1 percent per year, which is lower than its mean in the sample due to the presence of a very persistent inflation target process. As for monetary policy, it is fairly inertial, with $\rho_R$ around 0.7, and it exhibits a substantial degree of activism, with interest rates responding with a long-run coefficient of more than 2 to inflation and of almost 1 to output growth.

4. Potential Output and the Output Gap

With an estimated structural model in hand, we are now ready to explore the relationship between the actual economy, as observed over the past fifty years, and its unobserved efficient frontier, as inferred from the model. In our environment, the observed macroeconomic outcomes deviate from those under perfect competition, and thus from efficiency, due to the presence of monopoly power in goods and labor markets. This monopoly power, which stems from the imperfect substitutability of intermediate goods and of specialized labor services, allows firms to price their output above marginal cost and households to price their labor above the marginal rate of substitution.

In the aggregate, these markups create a wedge in the intratemporal efficiency condition, the equality of the marginal rates of substitution (MRS) and transformation (MRT) between consumption and leisure. More formally, define the aggregate price markup as

$$\mu_t^p = \frac{P_t}{MC_t},$$
Table 1. Prior distributions and posterior parameter estimates in the baseline model.

<table>
<thead>
<tr>
<th>Prior/Posterior</th>
<th>Dist</th>
<th>Mean</th>
<th>SE</th>
<th>Mode</th>
<th>Mean</th>
<th>SE</th>
<th>5%</th>
<th>Median</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital share N</td>
<td></td>
<td>0.30</td>
<td>0.05</td>
<td>0.16</td>
<td>0.16</td>
<td>0.01</td>
<td>0.15</td>
<td>0.16</td>
<td>0.17</td>
</tr>
<tr>
<td>Price indexation B</td>
<td></td>
<td>0.50</td>
<td>0.15</td>
<td>0.15</td>
<td>0.16</td>
<td>0.07</td>
<td>0.06</td>
<td>0.15</td>
<td>0.29</td>
</tr>
<tr>
<td>SS tech. growth N</td>
<td></td>
<td>0.50</td>
<td>0.03</td>
<td>0.47</td>
<td>0.47</td>
<td>0.02</td>
<td>0.43</td>
<td>0.47</td>
<td>0.51</td>
</tr>
<tr>
<td>Habit formation B</td>
<td></td>
<td>0.60</td>
<td>0.10</td>
<td>0.80</td>
<td>0.82</td>
<td>0.04</td>
<td>0.75</td>
<td>0.82</td>
<td>0.88</td>
</tr>
<tr>
<td>SS price markup N</td>
<td></td>
<td>0.15</td>
<td>0.05</td>
<td>0.25</td>
<td>0.25</td>
<td>0.04</td>
<td>0.18</td>
<td>0.25</td>
<td>0.32</td>
</tr>
<tr>
<td>SS log hours N</td>
<td></td>
<td>0.00</td>
<td>0.50</td>
<td>0.00</td>
<td>0.01</td>
<td>0.49</td>
<td>-0.79</td>
<td>0.00</td>
<td>0.80</td>
</tr>
<tr>
<td>SS inflation N</td>
<td></td>
<td>0.62</td>
<td>0.10</td>
<td>0.24</td>
<td>0.24</td>
<td>0.07</td>
<td>0.13</td>
<td>0.24</td>
<td>0.35</td>
</tr>
<tr>
<td>Discount factor G</td>
<td></td>
<td>0.25</td>
<td>0.10</td>
<td>0.17</td>
<td>0.18</td>
<td>0.05</td>
<td>0.10</td>
<td>0.18</td>
<td>0.27</td>
</tr>
<tr>
<td>Inverse Frisch G</td>
<td></td>
<td>2.00</td>
<td>0.75</td>
<td>2.36</td>
<td>2.73</td>
<td>0.76</td>
<td>1.62</td>
<td>2.67</td>
<td>4.12</td>
</tr>
<tr>
<td>Price stickiness B</td>
<td></td>
<td>0.66</td>
<td>0.10</td>
<td>0.84</td>
<td>0.84</td>
<td>0.03</td>
<td>0.79</td>
<td>0.84</td>
<td>0.89</td>
</tr>
<tr>
<td>Wage stickiness B</td>
<td></td>
<td>0.66</td>
<td>0.10</td>
<td>0.72</td>
<td>0.73</td>
<td>0.04</td>
<td>0.66</td>
<td>0.73</td>
<td>0.80</td>
</tr>
<tr>
<td>Elasticity util. cost G</td>
<td></td>
<td>5.00</td>
<td>1.00</td>
<td>5.10</td>
<td>5.31</td>
<td>1.01</td>
<td>3.77</td>
<td>5.25</td>
<td>7.08</td>
</tr>
<tr>
<td>Invest. adj. costs G</td>
<td></td>
<td>4.00</td>
<td>1.00</td>
<td>3.64</td>
<td>3.93</td>
<td>0.78</td>
<td>2.73</td>
<td>3.87</td>
<td>5.30</td>
</tr>
<tr>
<td>Reaction infl. N</td>
<td></td>
<td>1.70</td>
<td>0.30</td>
<td>2.35</td>
<td>2.32</td>
<td>0.21</td>
<td>1.97</td>
<td>2.33</td>
<td>2.66</td>
</tr>
<tr>
<td>Reaction GDP gr. N</td>
<td></td>
<td>0.40</td>
<td>0.30</td>
<td>0.82</td>
<td>0.85</td>
<td>0.14</td>
<td>0.63</td>
<td>0.84</td>
<td>1.10</td>
</tr>
<tr>
<td>Loading coeff. N</td>
<td></td>
<td>1.00</td>
<td>0.50</td>
<td>0.65</td>
<td>0.65</td>
<td>0.03</td>
<td>0.59</td>
<td>0.65</td>
<td>0.70</td>
</tr>
<tr>
<td>Auto. mp B</td>
<td></td>
<td>0.60</td>
<td>0.20</td>
<td>0.69</td>
<td>0.70</td>
<td>0.05</td>
<td>0.60</td>
<td>0.70</td>
<td>0.77</td>
</tr>
<tr>
<td>Auto. tech. B</td>
<td></td>
<td>0.40</td>
<td>0.20</td>
<td>0.13</td>
<td>0.13</td>
<td>0.06</td>
<td>0.04</td>
<td>0.13</td>
<td>0.23</td>
</tr>
<tr>
<td>Auto. gov. spending B</td>
<td></td>
<td>0.60</td>
<td>0.20</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Auto. investment B</td>
<td></td>
<td>0.60</td>
<td>0.20</td>
<td>0.69</td>
<td>0.69</td>
<td>0.05</td>
<td>0.60</td>
<td>0.69</td>
<td>0.77</td>
</tr>
<tr>
<td>Auto. price markup B</td>
<td></td>
<td>0.60</td>
<td>0.20</td>
<td>0.25</td>
<td>0.28</td>
<td>0.15</td>
<td>0.08</td>
<td>0.25</td>
<td>0.56</td>
</tr>
<tr>
<td>Auto. labor supply B</td>
<td></td>
<td>0.60</td>
<td>0.20</td>
<td>0.98</td>
<td>0.97</td>
<td>0.01</td>
<td>0.95</td>
<td>0.98</td>
<td>0.99</td>
</tr>
<tr>
<td>Auto. intertemporal B</td>
<td></td>
<td>0.60</td>
<td>0.20</td>
<td>0.68</td>
<td>0.63</td>
<td>0.08</td>
<td>0.49</td>
<td>0.64</td>
<td>0.76</td>
</tr>
<tr>
<td>Std mp IG1</td>
<td></td>
<td>0.15</td>
<td>1.00</td>
<td>0.22</td>
<td>0.22</td>
<td>0.01</td>
<td>0.20</td>
<td>0.22</td>
<td>0.25</td>
</tr>
<tr>
<td>Std tech. IG1</td>
<td></td>
<td>1.00</td>
<td>1.00</td>
<td>0.86</td>
<td>0.87</td>
<td>0.05</td>
<td>0.79</td>
<td>0.86</td>
<td>0.95</td>
</tr>
<tr>
<td>Std gov. spending IG1</td>
<td></td>
<td>0.50</td>
<td>1.00</td>
<td>0.36</td>
<td>0.36</td>
<td>0.02</td>
<td>0.34</td>
<td>0.36</td>
<td>0.40</td>
</tr>
<tr>
<td>Std investment IG1</td>
<td></td>
<td>0.50</td>
<td>1.00</td>
<td>6.98</td>
<td>7.56</td>
<td>1.46</td>
<td>5.39</td>
<td>7.45</td>
<td>10.17</td>
</tr>
<tr>
<td>Std price markup IG1</td>
<td></td>
<td>0.15</td>
<td>1.00</td>
<td>0.17</td>
<td>0.16</td>
<td>0.02</td>
<td>0.12</td>
<td>0.16</td>
<td>0.19</td>
</tr>
<tr>
<td>Std labor supply IG1</td>
<td></td>
<td>1.00</td>
<td>1.00</td>
<td>3.86</td>
<td>4.73</td>
<td>1.58</td>
<td>2.88</td>
<td>4.49</td>
<td>7.23</td>
</tr>
<tr>
<td>Std intertemporal IG1</td>
<td></td>
<td>0.10</td>
<td>1.00</td>
<td>0.03</td>
<td>0.03</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td>Std wage markup IG1</td>
<td></td>
<td>0.15</td>
<td>1.00</td>
<td>0.05</td>
<td>0.06</td>
<td>0.02</td>
<td>0.04</td>
<td>0.06</td>
<td>0.09</td>
</tr>
<tr>
<td>Std inflation target IG1</td>
<td></td>
<td>0.05</td>
<td>0.03</td>
<td>0.05</td>
<td>0.05</td>
<td>0.01</td>
<td>0.04</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>Std meas. error 1 IG1</td>
<td></td>
<td>0.15</td>
<td>1.00</td>
<td>0.49</td>
<td>0.49</td>
<td>0.03</td>
<td>0.45</td>
<td>0.49</td>
<td>0.55</td>
</tr>
<tr>
<td>Std meas. error 2 IG1</td>
<td></td>
<td>0.15</td>
<td>1.00</td>
<td>0.28</td>
<td>0.28</td>
<td>0.02</td>
<td>0.24</td>
<td>0.28</td>
<td>0.32</td>
</tr>
</tbody>
</table>
and the wage markup as
\[ \mu_t^w = \frac{W_t}{P_t MRS_t}, \]
so that
\[ MRS_t \mu_t^w \mu_t^p = MRT_t, \]
where we used the fact that nominal marginal cost \( MC_t = W_t/MPL_t \), and the marginal product of labor \( MPL_t \) is also the marginal rate of transformation between labor and final consumption (see also Gali, Gertler, and Lopez-Salido (2007)).

The equilibrium price and wage markups, \( \mu_t^p \) and \( \mu_t^w \), vary over time for two reasons. First, exogenous shifts in the substitutability of goods and labor services, and thus in the elasticity of their demand, affect firms and workers’ market power and desired markups. These shifts are captured by the stochastic processes \( \Lambda_{p,t} \) and \( \Lambda_{w,t} \), which correspond to the price and wage markup shocks, respectively, with standard deviations \( \sigma_p \) and \( \sigma_w \). The second source of equilibrium markup variation depends on the presence of nominal rigidities, which prevent firms and workers from achieving their desired markups at any given point in time. The resulting movements in average markups are endogenous and can be triggered by any shock hitting the economy.

Given these time-varying distortions, we define two counterfactual equilibria and associated levels of output.\(^5\)

**Definition 1.** **Efficient output.** In the limit for \( \xi_p \to 0, \xi_w \to 0, \sigma_p \to 0, \sigma_w \to 0, \lambda_p \to 0 \) and \( \lambda_w \to 0 \) (i.e. with flexible prices and wages, and perfectly substitutable intermediate goods and labor types) the baseline economy with monopolistic competition and sticky prices converges to a perfectly competitive economy, in which markups are zero. The First Welfare Theorem guarantees that the equilibrium of this economy is Pareto efficient. We call the equilibrium level of output in this counterfactual economy efficient output.\(^6\)

**Definition 2.** **Potential output.** In the limit for \( \xi_p \to 0, \xi_w \to 0, \sigma_p \to 0 \) and \( \sigma_w \to 0 \) (i.e. with flexible prices and wages, and constant elasticity of substitution among intermediate goods and labor types) the baseline economy with monopolistic competition and sticky prices

\(^5\) In the model, there is a small discrepancy between output and GDP, due to the presence of capital utilization costs. In the text, we sometimes refer to GDP as “output,” even if this usage is slightly imprecise.

\(^6\) This definition treats the exogenous path of government spending, \( G_t \), as part of the environment facing the planner, rather than as an instrument under its control.
converges to an economy with constant markups. We call the equilibrium level of output in this counterfactual economy potential output.

These definitions imply that, along the balanced growth path, actual and potential output grow at the same rate as efficient output. However, their level is inefficiently lower, due to the presence of the monopolistic distortions. Figure 1a plots the logarithm of U.S. GDP and the posterior median of potential output from the DSGE model. The latter is on average somewhat more stable than actual output, especially starting in the mid-eighties, although it is far from smooth.

To focus more closely on cyclical variation, figure 1b plots the posterior distribution of the model-implied output gap, which is the difference between output and the DSGE-based measure of potential. For comparison, 1b also plots the gap between actual and efficient output, in deviation from its steady state. Up to this level difference, the dynamics of the two gaps, and hence of potential and efficient output, are nearly identical. Therefore, in what follows, we ignore the small quantitative difference between these two constructs and treat the fluctuations in potential output as a measure of the movements in the model’s efficient frontier, which makes the output gap an indicator of the variation in the economy’s degree of inefficiency. The reason for focusing our measurement on potential, rather than on efficient GDP, is twofold. First, our model is mostly concerned with dynamics and has little to say about the level difference between potential and efficient output. Second, and most importantly, potential output is a more meaningful benchmark for optimal monetary policy, since monetary policy is neutral in the long-run, and therefore it cannot reduce the distance between actual and efficient output along the balanced growth path.

The shaded areas in figure 1 correspond to NBER recessions. They help highlight the pronounced cyclical behavior of the DSGE-based output gap, which peaks at the end of most expansions and declines during recessions. Looking at some specific episodes, figure 1b shows a declining output gap during the “twin” recessions of the early 1980s, which brought the economy back to its potential equilibrium after the notable “overheating” it experienced in the second half of the 1970s. This evolution of the DSGE gap is consistent with the idea

---

7 The log-linear dynamics of potential and efficient output are not exactly identical because the presence of fixed costs in the monopolistically competitive economy affects the elasticity of output with respect to changes in the inputs of production. In the competitive economy this effect is not present, since $\lambda_p \rightarrow 0$ also implies $F \rightarrow 0$, because the fixed cost is calibrated to offset the monopoly profits of the intermediate firms in steady state.
that monetary policy was at least partly responsible for loosing its grip on the real economy during the 1970s, and for bringing it back under control during Paul Volcker’s tenure as Fed Chairman. The estimates also indicate a slightly negative gap during most of the 1990s, implying that actual output was relatively slow in catching up with a surge in potential driven by the well-known pick-up in productivity growth. Finally, we observe that the last recession is associated with a substantial fall in output relative to the DSGE-based potential. This decline is driven by large negative investment shocks, which also contribute to depress potential output, and whose origin can be traced to financial disturbances further propagated by nominal rigidities, as argued in Justiniano, Primiceri, and Tambalotti (2011).
Overall, our assessment is that the DSGE output gap provides a plausible account of history, although one that is not necessarily in line with that of more traditional indicators of economic slack, such as detrended output. This discrepancy should not be surprising, however, since the DSGE gap is a measure of the distance of the economy from its efficient frontier, which is unlikely to evolve as smoothly as, say, trend output.

To quantify the relative role of efficient and inefficient fluctuations on business cycle dynamics, we decompose the deviation of GDP from a Hodrick-Prescott (HP) trend into two parts: the deviations of potential GDP and of the output gap from their respective HP trends. This is an exact decomposition, since the HP filter is linear and thus additive, but it is not orthogonal because potential output and the gap are correlated. Figure 2 shows the results of this exercise. First, note that the business cycle component of the DSGE gap peaks and troughs close to the NBER dates, with some delay in the last three “slow” recoveries. By eliminating the low-frequency component in the DSGE gap that is evident in figure 1b, the HP filter brings the fluctuations in the DSGE gap much more in line with the traditional chronology of expansions and recessions. Second, the decomposition highlights that the DSGE model’s estimate of potential output is roughly consistent with the Real Business Cycle view of fluctuations, in the sense that its efficient frontier displays cyclical movements of a similar magnitude to those of actual output (the correlation between HP-filtered potential and actual GDP is equal to 0.75).\footnote{Our neutral technology shock has a smaller impact on fluctuations than in traditional RBC models, due to the presence of investment adjustment costs and to a low estimate of the elasticity of labor supply. However, several other shocks also contribute to movements in the economy’s efficient frontier.} This volatility in the efficient frontier, however, coexists with a substantial volatility of the HP-filtered output gap, which is about 70 percent of that of detrended GDP (the correlation between the detrended gap and output is equal to 0.23). This evidence is the basis for our first important conclusion, that time variation in the economy’s degree of inefficiency is a crucial factor in macroeconomic fluctuations.

This conclusion is consistent with the evidence in Gali, Gertler, and Lopez-Salido (2007) and Sala, Soderstrom, and Trigari (2008), but stands in contrast to the results of many other studies in the empirical DSGE literature, as pointed out by Walsh (2005) and Mishkin (2007). This literature has typically found small DSGE output gaps, with little cyclical variation, but its findings can often be traced back to modeling assumptions that seem at odds with the data. For example, Andrés, López-Salido, and Nelson (2005) calibrate the coefficient of backward price indexation to one and do not include in their model markup shocks or
any other disturbance that directly affects the Phillips curve. With these tight restrictions, marginal costs are forced to explain the high frequency fluctuations in the first difference of inflation. To make this possible, the slope of the Phillips curve must be quite large, erroneously suggesting that nominal rigidities are nearly irrelevant for cyclical fluctuations. Edge, Kiley, and Laforte (2008) instead estimate a large scale model, with data starting in the early 1980s, but without an inflation target shock. Since the inflation rate exhibits a pronounced downward trend during this period, the inferred output gap inherits a similar trend. Finally, Levin, Onatski, Williams, and Williams (2005) detrend the real series prior to estimation, and assume that monetary policy responds directly to the output gap. Since their estimate of the output gap is small and countercyclical, this policy rule is at odds with the observed behavior of the Federal Reserve. If we impose the same restrictions assumed by these authors on our model, we can approximately replicate their results.
5. **Optimal Output and the (Lack of a) Policy Trade-Off**

The previous section documented sizable movements in the degree of inefficiency of the U.S. economy in the last fifty years. Why did stabilization policy not counteract these inefficient fluctuations? One possible reason is that these movements in the output gap represent the price policymakers had to pay to minimize other distortions. And indeed, in our economy, policymakers face a trade-off between output gap stabilization, on the one hand, and price and wage stabilization, on the other.

This trade-off stems from the fact that, in the equilibrium in which output is stabilized around potential, desired prices and wages—those that agents would set in the absence of nominal rigidities—are in general not constant. For example, desired prices change mechanically in response to changes in desired markups due to markup shocks. A more subtle reason for movements in desired prices is the coexistence of price and wage stickiness. With sticky wages, an increase in the marginal product of labor, due for example to an increase in productivity, will result in a fall in firms’ marginal costs, and hence in their desired price, since wages cannot fully adjust to absorb the productivity shock (Erceg, Henderson, and Levin (2000)).

Due to these movements in desired prices, workers and firms that reprice at different times will charge different prices. The resulting cross-sectional price and wage dispersion, whose movements over time go hand in hand with instability in price and wage inflation, is inefficient, since it forces workers and firms with identical tastes and technologies to supply different amounts of hours and of intermediate goods. Given that the aggregation of these inputs into aggregate utility and final output is concave, society would be better off if production and labor effort were shared equally across agents.

We can illustrate the inefficiency of an asymmetric distribution of the intermediate goods supply caused by price dispersion by aggregating the production functions (2.2) across firms $i$. Using demand (2.1), this aggregation produces a production function for the final good of the form

$$\Delta_t Y_t = A_t^{1-\alpha} K_t^\alpha L_t^{1-\alpha} - A_t F,$$

where

$$\Delta_t \equiv \int_0^1 \left( \frac{P_t(i)}{P_t} \right)^{1+A_{p,t}} P_{p,i} \, di$$
is a measure of price dispersion. Increases in price dispersion have the same effect as a fall in aggregate productivity, lowering the output of the final good, for any given level of the inputs. A similar dispersion term for wages directly reduces the utility of the average household, for any given level of the homogenous labor input $L_t$, due to the concavity of the labor aggregator (2.3).

To summarize, a stable output gap is in general incompatible with the absence of cross-sectional dispersion in prices and quantities in the intermediate goods and labor market, and therefore with stable price and wage inflation. As a result, stabilization policy faces a trade-off between these three objectives. In fact, in our economy, unlike in that of Erceg, Henderson, and Levin (2000) for instance, optimal policy needs to balance more than just these three goals, due to the presence of capital accumulation, and of all the other “frictions” needed to improve the model’s ability to fit the data.

For a more comprehensive and quantitative insight into these trade-offs, we turn to the study of the model’s optimal equilibrium.\footnote{Levin, Onatski, Williams, and Williams (2005), Schmitt-Grohé and Uribe (2007) and Christiano, Iliut, Motto, and Rostagno (2010) also compute optimal, or Ramsey, monetary policy in medium-scale DSGE models.} This corresponds to the equilibrium chosen by a planner that maximizes the utility of the average household under commitment, subject to the constraints represented by the behavior of private agents. The only instrument available to the planner is the short-term nominal interest rate, which defines this problem as one of optimal monetary policy. We characterize this optimal equilibrium following the approach developed by Benigno and Woodford (2006), which relies on a linear-quadratic approximation of the planner’s problem suitable for economies with an inefficient steady state. We then compute the path of output and of the other macro variables that would have been observed if policy had always been optimal, and the economy had been perturbed by the same sequence of shocks estimated in the baseline specification under the historical interest rate rule, except for $\pi_t^i$ and $\varepsilon_{R,t}$.\footnote{These two shocks do not affect the optimal equilibrium because they only appear in the interest rate rule (2.6), which is now replaced by optimal policy.}

Figure 3a compares actual, potential and optimal output. To do so in a parsimonious way, and to maintain comparability with figure 1 above, actual and optimal output are both presented in deviation from potential output. Figure 3b and c plot observed price and wage inflation, as well as their counterfactual evolution under the optimal policy. Optimal and potential output move extremely closely together, so that the optimal output gap—the
difference between optimal and potential output—is virtually zero at all times. The stability of the output gap under optimal policy is also consistent with a significant reduction in the volatility of price and, especially, wage inflation.

A large fraction of this reduction in inflation volatility is due to the fact that there is no time-varying inflation target $\pi^*_t$ under optimal policy. To isolate the contribution of the estimated variation in $\pi^*_t$ to inflation volatility, we compute the standard deviation of price inflation in a version of the baseline model with the estimated monetary policy rule, but with a constant inflation target. The standard deviation of quarterly inflation in this counterfactual simulation is 0.34, compared to 0.60 in the data and 0.26 under the optimal policy. This evidence suggests that the estimated movements in the inflation target, which are an integral part of the model’s description of historical Fed policy, are an important contributor to the suboptimality of observed inflation.

The optimality of a stable output gap is a well-known property of models with sticky prices and wages, as first shown by Erceg, Henderson, and Levin (2000). What is surprising about our findings is that a stable output gap is also consistent with roughly stable wage and price inflation. Hence the main result of the paper, that stabilization policy appears to have faced a negligible trade-off among its three main objectives over our sample. Much of the inefficient variation in output documented in section 4 could have been avoided, while at the same time reaping the benefit of more stable inflation. This result does not imply, however, that output would have been much smoother under optimal policy than under actual policy, since potential output also displays a significant amount of volatility, as shown in figure 2.

6. Wage Markup Shocks as the Main Source of the Trade-Off

The evidence in figure 3 suggests that a form of “divine coincidence”—in fact, a “trinity”—holds approximately in our estimated model. Price inflation, wage inflation and the output gap can be roughly stabilized at the same time. This result is all the more surprising, since the model includes three distinct sources of tension among these stabilization objectives: first, the rigidity in real wages, due to the simultaneous presence of wage and price stickiness; second and third, the exogenous variation in desired price and wage markups, due to shocks to market power. In this section, we show that among these three potential sources of a policy trade-off, wage markups are the one that matters most quantitatively. If the volatility
of desired wage markups is small, the trade-off is negligible; if the movements in desired markups are large, the trade-off becomes significant.

In our baseline specification, wage markup shocks are small, and the trade-off is negligible, for two main reasons. First, our approach to wage measurement suggests that most of the high frequency volatility in measured wages is due to noise, rather than to variation in workers’ market power. We illustrate the normative implications of this distinction in section 6.1. Second, we restrict a priori the low frequency movements in labor’s monopoly power. Since this modeling choice is controversial, in section 6.2 we also consider an extreme alternative specification, in which all exogenous shifts in labor supply are due to changes in desired wage markups. The surprising result of this last experiment is that the model’s optimal dynamics are very similar to the baseline, even in the presence of a significant trade-off.

6.1. High Frequency Wage Shocks: Markups or Noise? As just discussed, wage measurement is one of the two key factors responsible for the absence of a significant trade-off
between output gap and inflation stabilization in our estimated model. As described in section 3.1, we include two wage series in the measurement equation, following the general methodology proposed by Boivin and Giannoni (2006a). The idea behind this measurement approach is that each series contains useful information on the model’s concept of “the wage”, but that they both match this theoretical construct only imperfectly. Figure 4 illustrates the empirical underpinnings of this idea, by plotting the quarterly inflation rate in the two wage series we use. Compensation and earnings behave similarly at medium and low frequencies, but differ markedly at high frequencies, reflecting some conceptual differences in what they attempt to measure, and more in general the well-known difficulties in measuring aggregate wages (e.g. Abraham, Spletzer, and Stewart (1999), Bosworth and Perry (1994)). Given these large high-frequency discrepancies, the use of two wage indicators helps to distinguish idiosyncratic measurement factors—high frequency wage movements uncorrelated across the two indicators—from macroeconomic factors, which should produce fluctuations in both series.\footnote{In the estimation of our baseline model, we assume that the observation errors for the two wage series are uncorrelated. Relaxing this assumption would weaken identification, but would reinforce the substantive results of the paper, since even some correlated high-frequency movements in the two series might now be interpreted as observation errors, reducing the variation in wage markup shocks even further.}

To highlight the importance of this measurement choice, this section presents estimates of our model based on only one wage series with no measurement error, as in most empirical DSGE studies. For this exercise, we use the NIPA series of hourly compensation for the total economy, which is the more volatile of our two wage measures.\footnote{The typical estimation of DSGE models (for example Smets and Wouters (2007)) is conducted using compensation per hour in the non-farm business sector, which is even more volatile.} Table 2 reports the posterior estimates for this alternative empirical specification, which are generally very similar to those for the baseline. Two relevant exceptions are steady state inflation (which in the baseline is affected by the lower mean of the inflation rate in the compensation series), and the persistence of the price markup shock, which is now substantially higher.

The crucial difference in the estimates, however, is that the standard deviation of the wage markup shock is now almost six times larger than in the baseline. In addition, the variance decomposition of the model reveals that the explanatory power of this shock is large for wage inflation, especially at high frequencies, but very small for the real variables.\footnote{Wage markup shocks explain only 2 percent of GDP and hours volatility over the business cycle and 1 percent of consumption and investment.} \footnote{We have also estimated a model based on only one wage series with measurement error. The results are similar to our baseline, although identifying observation errors from wage markup shocks is more difficult in this model, since they have similar properties.}
These observations, combined with the dubious microfoundations of these shocks (Chari, Kehoe, and McGrattan (2009), Shimer (2009)), suggest that they are likely to reflect mostly observation error, rather than large variations in workers’ monopoly power from one quarter to the next.15

The extreme volatility of wage markup shocks in the model with only one wage measure, which we argued is empirically implausible, also has very misleading normative implications. This is because these large shocks directly affect workers’ desired wages, thus generating wide dispersion in relative wages when nominal adjustments are staggered. Given the large welfare costs due to wage dispersion, the planner is willing to pay a high price to reduce it as much as possible. Figure 5a shows that, in this model estimated with a single wage series, this price is paid in the form of extremely volatile optimal output, which is consistent with Rotemberg and Woodford (1999). To reduce the variability of wage and price inflation, as shown in

---

15 Results are similar when using only the earnings series, although the lower volatility of this wage measure implies somewhat smaller markup shocks than in table 2.
Table 2. Prior distributions and posterior parameter estimates in the model with only one observable wage series (compensation).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Type</th>
<th>Prior</th>
<th>Posterior</th>
<th>5%</th>
<th>Median</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>Capital share</td>
<td>N</td>
<td>0.30</td>
<td>0.17</td>
<td>0.01</td>
<td>0.16</td>
</tr>
<tr>
<td>( \tau_p )</td>
<td>Price indexation</td>
<td>B</td>
<td>0.50</td>
<td>0.06</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>( \tau_w )</td>
<td>Wage indexation</td>
<td>B</td>
<td>0.50</td>
<td>0.08</td>
<td>0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>( 100\gamma )</td>
<td>SS tech. growth</td>
<td>N</td>
<td>0.50</td>
<td>0.49</td>
<td>0.02</td>
<td>0.45</td>
</tr>
<tr>
<td>( h )</td>
<td>Habit formation</td>
<td>B</td>
<td>0.60</td>
<td>0.84</td>
<td>0.03</td>
<td>0.78</td>
</tr>
<tr>
<td>( \lambda_p )</td>
<td>SS price markup</td>
<td>N</td>
<td>0.15</td>
<td>0.29</td>
<td>0.04</td>
<td>0.23</td>
</tr>
<tr>
<td>( \log L_{ss} )</td>
<td>SS log hours</td>
<td>N</td>
<td>0.00</td>
<td>0.49</td>
<td>-0.76</td>
<td>0.04</td>
</tr>
<tr>
<td>( 100 (\beta^{-1} - 1) )</td>
<td>Discount factor</td>
<td>G</td>
<td>0.25</td>
<td>0.12</td>
<td>0.04</td>
<td>0.07</td>
</tr>
<tr>
<td>( \nu )</td>
<td>Inverse Frisch</td>
<td>G</td>
<td>2.00</td>
<td>1.96</td>
<td>0.69</td>
<td>1.32</td>
</tr>
<tr>
<td>( \xi_p )</td>
<td>Price stickiness</td>
<td>B</td>
<td>0.66</td>
<td>0.77</td>
<td>0.03</td>
<td>0.73</td>
</tr>
<tr>
<td>( \xi_w )</td>
<td>Wage stickiness</td>
<td>B</td>
<td>0.66</td>
<td>0.73</td>
<td>0.05</td>
<td>0.66</td>
</tr>
<tr>
<td>( \chi )</td>
<td>Elasticity util. cost</td>
<td>G</td>
<td>5.00</td>
<td>4.95</td>
<td>1.03</td>
<td>3.61</td>
</tr>
<tr>
<td>( S'' )</td>
<td>Invest. adj. costs</td>
<td>G</td>
<td>4.00</td>
<td>3.58</td>
<td>0.74</td>
<td>2.73</td>
</tr>
<tr>
<td>( \phi_\pi )</td>
<td>Reaction infl.</td>
<td>N</td>
<td>1.70</td>
<td>2.10</td>
<td>0.20</td>
<td>1.86</td>
</tr>
<tr>
<td>( \phi_N )</td>
<td>Reaction GDP gr.</td>
<td>N</td>
<td>0.40</td>
<td>0.94</td>
<td>0.13</td>
<td>0.73</td>
</tr>
<tr>
<td>( \rho_R )</td>
<td>Auto. mp</td>
<td>B</td>
<td>0.60</td>
<td>0.78</td>
<td>0.03</td>
<td>0.73</td>
</tr>
<tr>
<td>( \rho_z )</td>
<td>Auto. tech.</td>
<td>B</td>
<td>0.40</td>
<td>0.08</td>
<td>0.09</td>
<td>0.04</td>
</tr>
<tr>
<td>( \rho_g )</td>
<td>Auto. gov. spending</td>
<td>B</td>
<td>0.60</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>( \rho_\mu )</td>
<td>Auto. investment</td>
<td>B</td>
<td>0.60</td>
<td>0.74</td>
<td>0.05</td>
<td>0.63</td>
</tr>
<tr>
<td>( \rho_p )</td>
<td>Auto. price markup</td>
<td>B</td>
<td>0.60</td>
<td>0.90</td>
<td>0.05</td>
<td>0.78</td>
</tr>
<tr>
<td>( \rho_\varphi )</td>
<td>Auto. labor supply</td>
<td>B</td>
<td>0.60</td>
<td>0.98</td>
<td>0.01</td>
<td>0.96</td>
</tr>
<tr>
<td>( \rho_b )</td>
<td>Auto. intertemporal</td>
<td>B</td>
<td>0.60</td>
<td>0.44</td>
<td>0.11</td>
<td>0.28</td>
</tr>
<tr>
<td>( 100\sigma_R )</td>
<td>Std mp</td>
<td>IG1</td>
<td>0.15</td>
<td>0.21</td>
<td>0.01</td>
<td>0.20</td>
</tr>
<tr>
<td>( 100\sigma_z )</td>
<td>Std tech.</td>
<td>IG1</td>
<td>1.00</td>
<td>0.92</td>
<td>0.05</td>
<td>0.85</td>
</tr>
<tr>
<td>( 100\sigma_g )</td>
<td>Std gov. spending</td>
<td>IG1</td>
<td>0.50</td>
<td>0.37</td>
<td>0.02</td>
<td>0.34</td>
</tr>
<tr>
<td>( 100\sigma_\mu )</td>
<td>Std investment</td>
<td>IG1</td>
<td>0.50</td>
<td>0.70</td>
<td>1.37</td>
<td>5.66</td>
</tr>
<tr>
<td>( 100\sigma_p )</td>
<td>Std price markup</td>
<td>IG1</td>
<td>0.15</td>
<td>0.10</td>
<td>0.01</td>
<td>0.08</td>
</tr>
<tr>
<td>( 100\sigma_\varphi )</td>
<td>Std labor supply</td>
<td>IG1</td>
<td>1.00</td>
<td>2.99</td>
<td>1.09</td>
<td>2.11</td>
</tr>
<tr>
<td>( 100\sigma_b )</td>
<td>Std intertemporal</td>
<td>IG1</td>
<td>0.10</td>
<td>0.06</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>( 100\sigma_\omega )</td>
<td>Std wage markup</td>
<td>IG1</td>
<td>0.15</td>
<td>0.30</td>
<td>0.02</td>
<td>0.27</td>
</tr>
<tr>
<td>( 100\sigma_\pi )</td>
<td>Std inflation target</td>
<td>IG1</td>
<td>0.05</td>
<td>0.03</td>
<td>0.04</td>
<td>0.03</td>
</tr>
</tbody>
</table>
IS THERE A TRADE-OFF BETWEEN INFLATION AND OUTPUT STABILIZATION?

Figure 5. Actual and optimal GDP (in deviation from potential), price and wage inflation, in the model with only one observable wage series (compensation).

figures 5b and c, the planner must accept a path for optimal output that is significantly more volatile than that of actual output, at both business cycle and higher frequencies.

The striking implication of our calculations is that, in this case, optimal policy should have de-stabilized aggregate real activity, in order to reduce the volatility of price and, especially, wage inflation. In comparison, recall that this tension between output and inflation stabilization virtually disappears in the model with multiple wage indicators, in which the variation in wage markups is negligible.

Figure 6 confirms that the excess volatility of optimal output in this model is due to implausibly large wage markup shocks. The three panels plot GDP, price and wage inflation in the optimal allocation of an economy identical to the one depicted in Figure 5, but in which the wage markup shocks have been (arbitrarily) set to zero.\footnote{This experiment is very similar to one performed in Levin, Onatski, Williams, and Williams (2005) and reaches almost identical conclusions.} In this economy, the deviations
of optimal output from potential are much smaller than in figure 5, and more in line with those in the baseline model. The same is true for price and wage inflation, although the former is somewhat more volatile. This is because, in the one-wage model, the unconditional volatility of price markup shocks is higher than in the baseline estimation.

The conclusion we draw from these experiments is that ignoring measurement error in wages might lead to the erroneous normative implication that the output-inflation stabilization trade-off is substantial, and that, among their conflicting objectives, policymakers should prioritize the stabilization of wage inflation, even if this choice entails destabilizing output.

6.2. Low Frequency Labor Supply Shocks: Markups or Taste? As illustrated in section 2, the wage markup and labor disutility shocks are observationally equivalent in our empirical framework, and the only way to disentangle them is to posit a priori that they have different spectral profiles and hence contribute to fluctuations at different frequencies. This
is the approach we have followed, by assuming that the wage markup shock is \( i.i.d. \), while the labor disutility shock is a fairly persistent AR(1) process. This identification strategy captures the idea that the large low frequency variation in labor supply identified by the model is mostly attributable to shifts in households’ preferences for market work, such as the secular change in attitudes towards female labor participation that started in the 1960s, or other slow moving demographic developments. However, if we were to attribute all or part of this low frequency variation in hours to changes in the competitiveness of the labor market, our estimate of potential output would change significantly, since this shock directly affects the distance of the economy from its efficient frontier.

Figure 7a illustrates this point quantitatively, by comparing the estimates of potential output arising from two alternative interpretations of the source of low frequency fluctuations in hours, as in Sala, Söderström, and Trigari (2010). The first interpretation corresponds to our baseline assumption that these low frequency movements are mainly due to labor disutility shocks. Under the alternative interpretation, there are no taste shocks, and the low-frequency fluctuations in hours originate from changes in the monopoly power of workers, which do not affect the potential equilibrium.\(^\text{17}\) The solid line in figure 7a represents the difference between the estimates of potential output under the two scenarios, baseline and “all-wage-markup.” This difference is very large—ranging from \(-10\) to \(+15\) percent—but most of its variation is at low frequencies. This should not be too surprising, given that the shifts in the disutility of labor we estimate must be highly persistent and very large to explain the wide secular swings in total hours observed in the sample.

We just showed that the behavior of potential output changes considerably with respect to our baseline estimates under the extreme assumption that labor supply shocks only reflect changes in workers’ market power. However, this is not the case for optimal output, as seen from the dashed line in figure 7a, which depicts the difference between the estimates of optimal output under the two alternative scenarios. This difference is an order of magnitude smaller than that between the estimates of potential. Optimal output is relatively invariant to the source of labor supply shifts because the planner has a strong distaste for wage dispersion, and thus chooses a very stable path for wage inflation, as in the baseline case of section 5.

\(^{17}\) Given the observational equivalence of these two alternative assumptions, the model does not need to be re-estimated.
Figure 7. Effect of alternative assumptions on the origin of labor supply shocks on potential and optimal GDP. In panel (a), the blue line represents the difference between potential GDP in the all-wage-markup specification and the baseline. The red dashed line is the difference in optimal GDP across the same two models. Panel (b) plots the difference between optimal and potential GDP in the two specifications.

As a result, quantities—hours and output in particular—bear the brunt of the adjustment to labor supply shocks in the optimal equilibrium, regardless of these shocks’ origin.

Figure 7a implies that the optimal output gap—optimal minus potential—is subject to large swings in the all-wage-markup model, unlike in the baseline. This is shown in panel (b), and reflects the fact that optimal output is very similar in the two specifications, while potential output is not. The volatility of the output gap in the optimal equilibrium of the all-wage-markup model reflects a significant policy trade-off, which contrasts with the lack of such a trade-off that characterizes the baseline specification. However, most of the variation in the output gap is concentrated at low frequencies, and is caused by changes in market power,
which suggests that monetary policy might not be the most appropriate tool to address this
distortion, as compared to more targeted fiscal or microeconomic policies.

Moreover, the fact that optimal output is roughly invariant in the two scenarios considered
in figure 7 implies that researchers interested in characterizing the equilibrium implications
of optimal policy do not need to take a strong stance on the ultimate sources of labor supply
shocks, an issue that remains unsettled in the literature (e.g. Gali, Smets, and Wouters
(2011) and Christiano, Trabandt, and Walentin (2010)). One consequence of this finding is
that medium scale New-Keynesian models like the one presented here might be more useful
for monetary policy analysis than suggested for instance by Chari, Kehoe, and McGrattan
(2009).

7. Conclusions

We estimated a DSGE model of the U.S. economy. In this model, workers and firms
have some monopoly power, and markups of price over marginal cost and of wages over the
marginal rate of substitution fluctuate due to exogenous changes in monopoly power, as well as
to the presence of sticky prices and wages. According to our estimates, the movements in these
markups were associated with large fluctuations in U.S. GDP away from its efficient frontier.
Moreover, we find that monetary policy could have virtually eliminated these inefficient
output movements, reducing at the same time the volatility of price and wage inflation, and
thus achieving an improvement in welfare for the model’s representative agent.

The key factor driving this surprising absence of a meaningful trade-off between output
(gap) and inflation stabilization in our model is that exogenous variation in desired wage
markups contribute little to fluctuations, at both high and low frequencies. At high frequen-
cies, we infer small markup shocks due to our approach to wage measurement, which consists
of matching the model’s “wage” to two measures of the return to labor. As a result, we are
better able to isolate high frequency idiosyncrasies specific to each series, from a common
component that is more likely to represent genuine macroeconomic forces. Indeed, the same
model estimated with only one wage series, as in most of the literature, finds implausibly
large high frequency fluctuations in desired wage markups, which generate a strong tension
between output gap and inflation stabilization. In this alternative specification, optimal pol-
icy would be forced to resolve this conflict by inducing higher output volatility than observed
in the data, in order to obtain a smoother path of inflation.
At low frequencies, wage markup shocks are small due to our prior, which assumes that secular shifts in labor supply are more likely to derive from changes in households’ taste for market work, than from movements in the monopoly power of workers. When we assume the opposite, inefficient output fluctuations continue to be large, but they also become the source of a costly trade-off for monetary policy. However, this increase in the tension between the conflicting stabilization objectives of the central bank has only a minor impact on its optimal behavior, and hence on the behavior of the economy under optimal policy.

Our results point to a significant discrepancy between the model’s equilibrium under historical monetary policy, as described by the estimated interest rate rule, and the optimal equilibrium. What are the reasons for this discrepancy? One possibility is that the model’s welfare function is not a good representation of the actual objectives of U.S. monetary policy. For example, wage distortions loom large in the utility of the representative agent in the model, but they are seldom mentioned as a direct preoccupation by policymakers. A second possibility is that we might have overlooked some relevant constraint in setting interest rates. The most obvious one in this respect is the zero lower bound on nominal interest rates. However, a preliminary exploration of this issue suggests that this bound is not violated often in the optimal equilibrium. Finally, actual monetary policy might have been misguided, at least at some points in the past. This is a fairly common conclusion among researchers, especially for the period roughly between 1965 and 1980 (e.g., Clarida, Gali, and Gertler (2000), Cogley and Sargent (2005) and Primiceri (2006)). The fact that we assume an invariant policy rule across the entire span of our long sample prevents us from addressing this possibility with finer historical detail, although doing so would be useful.

Appendix A. Normalization of the Shocks

As in Smets and Wouters (2007), some of the exogenous shocks are re-normalized by a constant term. In particular, we normalize the price and wage markups shocks and the intertemporal preference shock.

More specifically, the log-linearized Phillips curve is

\[ \dot{\pi}_t = \frac{\beta}{1 + \beta_t} E_t \dot{\pi}_{t+1} + \frac{1}{1 + \beta_t} \pi_{t-1} + \kappa \dot{s}_t + \kappa \dot{\lambda}_{p,t}. \]

The normalization consists of defining a new exogenous variable, \( \dot{\lambda}_{p,t}^* \equiv \kappa \dot{\lambda}_{p,t} \), and estimating the standard deviation of the innovation to \( \dot{\lambda}_{p,t}^* \) instead of \( \dot{\lambda}_{p,t} \). We do the same for the wage
IS THERE A TRADE-OFF BETWEEN INFLATION AND OUTPUT STABILIZATION? 35

markup and the intertemporal preference shock, for which the normalizations are

\[
\hat{\lambda}_{w,t}^* = \left( \frac{(1 - \beta \xi_w)(1 - \xi_w)}{(1 + \nu(1 + \frac{1}{\xi_w})) (1 + \beta) \xi_w} \right) \hat{\lambda}_{w,t}
\]

\[
\hat{b}_t^* = \left( \frac{(1 - \rho_h) (e^\gamma - h \beta \rho_h) (e^\gamma - h)}{e^{\gamma h} + e^{2 \gamma} + \beta h^2} \right) \hat{b}_t.
\]

These normalizations are chosen so that these shocks enter their equations with a coefficient of one. In this way, it is easier to choose a reasonable prior for their standard deviation. Moreover, the normalization is a practical way to impose correlated priors across coefficients, which is desirable in some cases. For instance, imposing a prior on the standard deviation of the innovation to \( \lambda_{p,t}^* \) corresponds to imposing priors that allow for correlation between \( \kappa \) and the standard deviation of the innovations to \( \lambda_{p,t} \). Often, these normalizations improve the convergence properties of the MCMC algorithm.

REFERENCES


