Re-estimating Potential GDP: New Evidence on Output Hysteresis*

Diego Anzoategui — Rutgers University

Min Kim — Rutgers University

August 24, 2021

Click Here for Latest Version

Abstract

We propose a simple structural method to estimate potential GDP. Our approach is derived from a standard New Keynesian model, yet it is consistent with a wide range of structural assumptions. Moreover, it is not subject to the Lucas Critique, it does not resort to Bayesian estimation of the underlying model, and it is consistent with a large set of possible parametrizations. We estimate potential GDP for the US and use our series to contribute to the debate on the effects of demand shocks on aggregate supply. We find evidence supporting hysteresis hypotheses claiming that demand shocks can affect potential GDP.

Keywords: Potential GDP, Business Cycle, Hysteresis.

JEL Classification: E30, E32, E37

*For comments and discussions, we thank seminar participants at Rutgers Macro Seminar, the 2021 Annual Conference of the Canadian Economics Association, the Third Warsaw Money-Macro-Finance conference, and EcoMod 2021.

Diego Anzoategui (corresponding author): diego.anzoategui@rutgers.edu, Min Kim: mk1564@economics.rutgers.edu.

1 Introduction

Potential or flexible-price GDP is the level of output in the absence of price or wage rigidities. This indicator is extensively used for two important purposes. First, economists use potential GDP as a predictor of actual GDP in the long run. The reason is that prices become more flexible at longer time horizons and, therefore, GDP and potential GDP tend to be closely related at low frequencies. Second, potential GDP is a useful indicator for monetary policy design. The difference between GDP and potential GDP, a measure called output gap, is an important factor affecting inflation. Positive output gaps are related with inflationary pressures, whereas negative ones tend to reduce the inflation rate. In fact, monetary theory implies that, absent mark-up shocks, monetary authorities should aim to a null output gap to maximize welfare.

Due to its evident importance for economic policy, many institutions construct their own potential GDP series: the Congressional Budget Office (CBO), Federal Reserve, International Monetary Fund, OECD, among others. However, there is not clear consensus on the best estimation method. Potential GDP is a counterfactual and, thus, its estimation requires specific assumptions. In fact, different methods have their own strengths and weaknesses. One can classify the available methods into two broad sets: non-structural methods and dynamic stochastic general equilibrium (DSGE) approaches.[1]

Methods that are not based on structural models are either purely statistical calculations of the low-frequency component of GDP, or estimates that involve the use of reduced-form equations to infer the flexible-price GDP counterfactual.[2] They have the advantage of being generally straightforward because they do not require the estimation of a structural model and involve simple computation. Nevertheless, they suffer from important shortcomings. First, they generally depart from the relevant economic theory and, as a result, do not estimate the flexible-price output level providing little guidance for the design of monetary policy. Second, they are subject to the Lucas (1976) critique because they use reduced-form regressions to infer counterfactual levels of GDP. As a result, potential output estimates coming from these methods might be biased.

[1]See Mishkin (2007) for brief description of the methods. Coibion et al. (2017b) details the approaches used by several organizations.

[2]This includes methods that compute trends such as the Hodrick and Prescott (1981) filter or methods related to Beveridge and Nelson (1981). Methods that employ reduced form equations such as the Okun's Law or the Phillips Curve are included in this set too.

On the other hand, potential GDP estimation using a DSGE model provides an unbiased estimate of the flexible-price output level and is not subject to the Lucas critique. Hence, this approach tends to be more informative for economic policy design. However, this advantage comes with shortcomings too. DSGE-based methods require the estimation of a generally large set of structural parameters and shock processes. As a consequence, the typical concerns regarding parameter identification in these models apply.[3] Furthermore, the results are model-specific. In fact, applications of this approach seem to provide divergent predictions. For instance, Andrés et al. (2005), Neiss and Nelson (2005), and Edge et al. (2008) use DSGE models to estimate potential GDP and find small and less persistent output gaps compared to the ones computed using conventional detrending methods. On the other hand, Sala et al. (2008), Galí et al. (2012) and Justiniano et al. (2013) find that DSGE-based output gaps are similar to non-structural estimates.

This paper proposes a new DSGE-based method to compute potential GDP that combines the strengths of both approaches. First, it is as simple as any non-structural method. Second, it provides an unbiased estimate of flexible-price output and it is not subject to the Lucas Critique because it is based on a structural model. Third, unlike other model-based approaches, our method does not require prior knowledge of all deep parameters not does it resort to Bayesian estimation or calibration of the underlying model. Hence, the estimates are consistent with a large set of possible parametrizations.

Our baseline method is derived from a textbook version of the New Keynesian model with wage rigidities and without endogenous capital stock. However, the method is fairly general; it remains unchanged to a wide set of modifications to structural assumptions related to preferences, production technology, monetary policy rules, and expectation formation. It is marginally modified if we incorporate price rigidities, government spending shocks, or even a labor supply shock to capture the supply side impact of lock-down measures related to the COVID-19 pandemic. Moreover, our method can precisely estimate potential GDP series generated by more complicated models such as Smets and Wouters (2007), especially when short-run labor wealth effects are unimportant.

We apply our methodology to estimate potential GDP in the US and find that our estimate is highly correlated with the one computed by the CBO, a widely used measure of potential GDP.

[3] See, for example, Canova and Sala (2009).

However, we do find a stark difference between the two series during and after the Great Recession. Our estimation points to an increase in potential output during the recession in 2009 but a poor growth rate from 2010 onward. In turn, CBO's series picture a decline in potential GDP growth during the recession but a much higher average growth rate after 2010. These results are robust to different estimation specifications.

Lastly, using our estimated potential GDP series, we contribute to an active debate on the effects of demand shocks on GDP in the medium or long run. We analyze the impact of exogenous demand shocks on potential GDP and find evidence supporting the hysteresis hypothesis; typical demand shocks like monetary or fiscal shocks have a significant impact on potential output, explaining their persistent impact on GDP.

Related Literature. This paper is directly related to contributions using DSGE models to estimate potential GDP such as Andrés et al. (2005), Neiss and Nelson (2005), Edge et al. (2008), Sala et al. (2008), Basu and Fernald (2009), Galí et al. (2012) and Justiniano et al. (2013). In order to compute potential GDP, these papers calibrate or estimate a large set of deep parameters and strucutral shock processes. As mentioned above, our paper proposes a much simpler approach where parameter and structural shock identification is clearer. In fact, prior knowledge of all deep parameters is not required.

Moreover, this paper is similar to Coibion et al. (2017b) in spirit. Coibion et al. (2017b) highlights shortcomings in existing estimation approaches employed by several institutions and propose a non-model-based method following the contribution of Blanchard and Quah (1989). In a similar way, Gordon (2014) proposes univariate methods to compute potential GDP. Differently from these contributions, we present a method that is DSGE-based and, as a result, provides an unbiased estimate of the flexible-price level of output.

Our baseline method consists of estimating a Structural Vector Autoregression (SVAR) to obtain a counterfactual level of GDP that is immune to the Lucas Critique. In that sense, our contribution is related to Beraja (2019) who offers a method to compute counterfactuals using SVARs. Compared with Beraja (2019), our method is more specific as it is only useful to compute potential GDP, but it has the advantage of being simpler because it requires fewer steps and, importantly, no prior knowledge of policy rules parameters is needed. In deriving the method, we benefited from several papers linking DSGE models with their SVAR representations. In particular, we used directly or indirectly insights from Fernández-Villaverde et al. (2007), Ravenna (2007) and Beraja (2019),

Lastly this paper is related to a sequence of contributions empirically analyzing the presence of hysteresis.[4] One of the first papers providing evidence on permanent impacts of temporary exogenous shocks is Cerra and Saxena (2008). More recently, Blanchard et al. (2015) show that most of the economic recessions suffer from hysteresis. Jordà et al. (2013) and Reinhart and Rogoff (2014) find that financial crises generate recessions that are more persistent. By extending the method of Blanchard and Quah (1989), Furlanetto et al. (2020) identify demand shocks with long-run effects on output. Analyzing specific shocks, Fatás and Summers (2018) and Jordà et al. (2020) show evidence of permanent effects of fiscal and monetary policy, respectively. We contribute to this literature by showing the effects of demand shocks on potential output, a channel that helps to explain the persistent effects on actual GDP.

Outline. This paper continues as follows. Section 2 provides a broad picture of the empirical strategy to infer potential GDP using DSGE models. Section 3 describes the baseline method proposed in this paper, section 4 applies our method to the US data, whereas section 5 discusses extensions to our method. Section 6 analyzes the responses of our potential GDP estimates to demand shocks. Based on our previous results, section 7 discusses estimation issues that emerge when we allow for endogenous TFP in our method. Section 8 concludes.

2 DSGE-based Potential GDP Estimation: A Big Picture

New Keynesian (NK) models like the ones proposed by Christiano et al. (2005) or Smets and Wouters (2007) are basically Real Business Cycle models with an endogenous labor wedge, as defined by Chari et al. (2007).[5] The log of the labor wedge is represented by the log difference between the marginal product of labor and the marginal rate of substitution between consumption and leisure. In other words, this wedge summarizes distortions that make the social marginal benefit from working different from its marginal cost. In NK models, omitting markup shocks,

[4]See Cerra et al. (2020) for a complete literature review.

[5]The two important features that generate an endogenous labor wedge are: (i) monopolistic competition in either goods or labor markets, and (ii) nominal wage or price rigidities.

cyclical fluctuations in the labor wedge are directly linked with nominal rigidities. In fact, following Galí et al. (2007) the labor wedge is equal to the sum of the wage markup and the price markup (in log-deviations from steady state values). That is,

$$\underbrace{\tau_t}_{\text{labor wedge}} = \underbrace{\mu_t^w}_{\text{wage markup}} + \underbrace{\mu_t^p}_{\text{price markup}}$$

where the wage mark up μ_t^w is defined by the difference between the real wage and the marginal rate of substitution between consumption and leisure, and the price markup μ_t^p is the difference between the marginal product of labor and the real wage. Both μ_t^w and μ_t^p are endogenous objects that fluctuate in the business cycle and summarize the propagation of shocks related to both nominal wage and price rigidities, respectively. Hence, a basic description of how GDP is affected by shocks in NK models can be summarized by the following diagram.

$$\begin{tikzcd} \text{Shocks} \arrow[r, "\text{Nominal rigidities}"] \arrow[rr, bend right, "\text{Other propagation mechanisms}"'] & \tau_t \arrow[r] & y_t \text{ (gdp)} \end{tikzcd}$$

For a given set of states, structural shocks affect GDP through two broadly defined channels. First, the presence of nominal rigidities generates fluctuations in the labor wedge, affecting the level of GDP. This first channel is summarized by the red arrows in the diagram above. Second, the blue arrow highlights other shock propagation mechanisms on GDP; these are generally related to real rigidities such as adjustment costs, habit formation, etc.

The previous discussion is useful to describe a general strategy to obtain potential output from a NK model. Potential GDP is the counterfactual level of output under flexible prices. Hence, NK models calculate this counterfactual by shutting down the propagation mechanism linked to nominal rigidities (the red arrows in the diagram above). Hence, potential GDP is the level of output under the counterfactual assumption of a fixed labor wedge. In fact, the method that we describe in the next section is just a simple way of eliminating labor wedge fluctuations that is immune to the Lucas Critique.

Before getting into a detailed description of our approach, a couple of general comments on the method are in order. Potential GDP series derived from our method exclude the effect of markup shocks, which are exogenous changes in wage and price markups. The reason for this exclusion is twofold. First, these shocks do not have a clear structural interpretation as noted by Chari et al. (2009). Second, their importance in early estimations of NK models like Smets and Wouters (2007) has to do mostly with high measurement error in real wage data series. In fact, once measurement error is allowed in the estimation of these models, results point to a reduced role of these shocks. Hence, including or excluding them from the definition of potential GDP seems quantitatively irrelevant.[6]

3 A Simple Method to Estimate Potential GDP

This section describes a simple method to estimate potential output. We first describe the simplest version of our method, defined as the "baseline" method. This estimation strategy is derived from a textbook version of the NK model as in Woodford (2003) or Galí (2015) with sticky wages following Erceg et al. (2000). We then show that this simple method is much more general than it looks, as it is consistent with many other macroeconomic models with different technologies, preferences, monetary policy rules and expectation formation assumptions. Lastly, we assess the quantitative performance of our method in short samples.

3.1 The Baseline Model

The economy is populated by a representative household, a representative final good producer, intermediate good producers, labor unions and a monetary authority.[7] The household's preference is described by a log-utility function with consumption habit formation and disutility of labor. Its consumption-savings decision is summarized by the following (log-linearized) Euler equation,

$$-i_t = -mu_t + \mathbb{E}_t mu_{t+1} - \mathbb{E}_t \pi_{t+1} + \sigma_z \left(\mathbb{E}_t z_{t+1} - z_t \right) \tag{1}$$

[6] See for example Justiniano et al. (2013).

[7] Note that all non-stationary variables are detrended by their growth rates at the Balanced Growth Path in this section. Given that the model is standard, we provide key log-linearized equations in this section. We describe the model details in an Online Appendix.

where i_t denotes the nominal interest rate, π_t is the net inflation rate, and $\sigma_z z_t$ is an iid preference shock normally distributed with standard deviation σ_z. Finally, mu_t represents the marginal utility of consumption, defined as follows,

$$mu_t \equiv -\frac{1+g_y}{1+g_y-h}c_t + \frac{h}{1+g_y-h}c_{t-1} \tag{2}$$

where c_t is (detrended, log-linearized) consumption, h is the habits parameter, and g_y is the growth rate of output (and consumption) at the balanced growth path (BGP).

Intermediate good producers have access to a Cobb-Douglas production function in labor (for simplicity, there is no capital in the baseline model) defined by

$$y_t = a_t + (1-\alpha)n_t \tag{3}$$

where y_t represents output, n_t is labor (total hours), and a_t denotes Total Factor Productivity (TFP) that follows a random walk,

$$a_t = a_{t-1} + \sigma_a \varepsilon_{at} \tag{4}$$

where ε_{at} is standard normally distributed and $\sigma_a > 0$.[8] The behavior of intermediate good producers can be summarized by their pricing decision. Given that there are no price rigidities in the baseline model, firms charge a fixed markup over nominal marginal costs. This implies that real marginal cost is constant and equal to its steady-state value (in log-deviations, $mc_t = 0$), meaning that the following expression holds,

$$w_t - \frac{1}{1-\alpha}a_t + \frac{\alpha}{1-\alpha}y_t = mc_t = 0 \tag{5}$$

where w_t denotes the real wage. Labor unions hire labor from the household and provide labor services to firms. These unions face a Calvo (1983) style friction when deciding to determine the nominal wages they charge to firms. The maximizing behavior of these unions imply the following

[8]Our assumption of a random walk follows empirical evidence. Estimates of TFP for the US such as Fernald (2012) indicate the presence of a unit root.

wage Phillips curve,

$$\pi_t^w = -\kappa^w \mu_t^w + \beta \mathbb{E}_t \pi_{t+1}^w$$

where π_t^w is the nominal wage inflation, κ^w is the slope of the Phillips curve, $\beta \in (0, 1)$ is the discount factor, and μ_t^w represents the wage markup defined as in Galí et al. (2007) by

$$\mu_t^w \equiv w_t - \varphi n_t + mu_t \tag{6}$$

where φ is the inverse Frisch elasticity. Finally, the goods market clearing condition and monetary policy rule are decribed by equations (7) and (8) below,

$$y_t = c_t \tag{7}$$

$$i_t = \phi_\pi \pi_t + \sigma_i \nu_t \tag{8}$$

where $\sigma_i \nu_t$ is a normally distributed iid monetary policy shock with standard deviation σ_i, and $\phi_\pi > 1$. Equations (1) to (8) provide a complete description of our framework.

3.2 The Baseline Method

What is the model-implied potential GDP? As explained above, potential GDP is the level of GDP that the economy would have if there were no nominal rigidities. Hence, in the baseline model, potential output is the flexible-wage counterfactual. We can actually compute that output level using the fact that wage markups are constant at their steady-state levels under flexible wages, or $\mu_t^w = 0$. Applying the constraint $\mu_t^w = 0$ to (6) and using (3), (4) and (5), we get an expression describing potential output growth Δy_t^p in this model,

$$\Delta y_t^p = \theta_1 \Delta y_{t-1}^p + \theta_0 \varepsilon_{at} \tag{9}$$

where,

$$\theta_0 \equiv \frac{\frac{1+\varphi}{1-\alpha}\sigma_a}{\frac{1+g_y}{1+g_y-h} + \frac{\alpha+\varphi}{1-\alpha}} \qquad \text{and} \qquad \theta_1 \equiv \frac{\frac{h}{1+g_y-h}}{\frac{1+g_y}{1+g_y-h} + \frac{\alpha+\varphi}{1-\alpha}}$$

Equation (9) reveals that in order to get potential GDP estimates we need two things: (i) first, estimates of parameters θ_0 and θ_1 which are in turn functions of deep parameters of the model, and (ii) identified productivity shocks ε_{at} consistent with the model. Below, we propose a simple method to perform both tasks without estimating or calibrating the whole set of parameters or structural shocks. This simple method is described in proposition 1.

Proposition 1. *θ_0, θ_1 and ε_{at} in equation (9) can be estimated from the following Structural Vector Autoregression (SVAR) estimation,*

$$\begin{bmatrix} \Delta y_t \\ \mu_t^w \end{bmatrix} = \mathbf{B} \begin{bmatrix} \Delta y_{t-1} \\ \mu_{t-1}^w \end{bmatrix} + \mathbf{C} \begin{bmatrix} \varepsilon_{at} \\ \xi_t \end{bmatrix} \tag{10}$$

where Δy_t is the GDP growth rate and ξ_t is a weighted average of demand shocks. In particular, letting c_{ij} and b_{ij} for $i, j = \{1, 2\}$ be the elements of the 2×2 matrices $\mathbf{B}$ and $\mathbf{C}$,

$$\theta_0 = c_{11} - \frac{c_{21}c_{12}}{c_{22}} \qquad \theta_1 = b_{11} - \frac{b_{21}c_{12}}{c_{22}}$$

And ε_{at} can be calculated using forecast errors and $\mathbf{C}$.

Proof. See Appendix A. □

Proposition 1 states that it is possible to estimate potential GDP by simply estimating a SVAR. This is a much simpler process than a full-model estimation using Bayesian methods, and the identification of parameters and shocks is clearer too. The proposition shows that there is no need to estimate all the deep parameters of the model. In turn, it suffices to get estimates of certain functions of those parameters, θ_0 and θ_1. Moreover, Proposition 1 clearly determines what a researcher needs to carry out the estimation: (i) series of GDP growth (Δy_t) and wage markups (μ_t^w), and (ii) an identification strategy to estimate matrix $\mathbf{C}$.

The choice of the wage markup series and the strategy to estimate $\mathbf{C}$ are far from obvious. However, the labor wedge literature and contributions on the estimation of SVARs provide clear guidance in this regard. In terms of measuring the wage markup, we follow Galí et al. (2007) in our baseline case and assume a log utility with a Firsch elasticity equal to one. Notice that it is only in this stage that we need to claim knowledge of a deep parameter of the model: we need to assume a value for the Frisch elasticity. Nevertheless, it is straightforward to perform robustness analysis using a set of values for the parameter and comparing outcomes.[9]

In terms of the matrix $\mathbf{C}$ estimation, contributions on SVAR identification suggest two broad strategies. First, taking advantage that TFP has a unit root in the model, we could impose long-run restrictions as in Blanchard and Quah (1989) to identify matrix $\mathbf{C}$. Second, following Stock (2008), we could use exogenous shocks identified in other papers as instrumental variables to identify $\mathbf{C}$. Our preferred method is the SVAR-IV proposed by Stock (2008) because, in our empirical application for the US, we find that long-run restrictions are not robust and tend to be severely affected by the sample chosen for estimation.[10]

3.3 How general is the Baseline Method?

In the previous sections, we derived a method to estimate potential GDP from a simple NK model with specific assumptions. However, it is straightforward to show that our potential GDP estimation method is fairly general and remains unchanged to several modifications to the baseline model. Proposition 2 describes how general the method is.

Proposition 2. *As long as they do not add state variables, the following modifications to the baseline model do not change the method proposed in proposition 1:*

1. *Changing the production function to other production functions that only use labor as input*
2. *Using other preferences consistent with a BGP, with or without habit formation*
3. *Adding wage markup shocks*

[9] See the Online Appendix.

[10] Results available upon request. As noted by Stock (2008), this is a common concern that can be interpreted as a weak instrument problem.

4. *Assuming other monetary policy rules*
5. *Other expectation formation assumptions*
6. *Adding capital utilization as in Smets and Wouters (2007) and assuming capital in fixed supply (growing in a BGP)*

Proof. See Appendix B. □

Proposition 2 states that the baseline method admits changes in assumptions related to technology, preferences, monetary policy rule and expectation formation assumptions. Moreover, it is possible to add more shocks like wage markup shocks or even add capital utilization in a standard way. The key requirement to keep the method unchanged is that the set of state variables of the model must remain unchanged. This is because any new state variable incorporated after a modification must be included as a right hand side variable in the system (10), changing the empirical strategy.

Note, however, that there could be many other modifications (not included in proposition 2) that modify the baseline method but marginally. This is, for example, the case of adding typical assumptions in medium scale NK models such as wage indexation and nominal interest rate smoothing. Under these two assumptions, the model has two additional state variables (lagged wage inflation and nominal interest rate) that would need to be incorporated into system (10) without changing the derivation of potential GDP from the system coefficients.

There are of course obvious additions to the baseline model that will change the method more significantly like adding price rigidities. This is a potentially important extension that we will consider in section 5 for robustness.

3.4 Method performance in small samples

Proposition 1 clearly defines a strategy to estimate potential GDP. However, given the relative lack of long series of macroeconomic variables, it is important to analyze the performance of the method in short samples. With that aim, we test our method with simulated data coming from our baseline model.

We first calibrate the model using standard parameters in the literature as shown in table 1. The calibration of structural shock processes deserves some comments. The shocks are assumed to be iid and their standard deviations are calibrated as follows. The standard deviation of TFP shocks matches that of Fernald (2012) TFP series, whereas the volatility of preference and monetary shocks are set to match the standard deviation of wage markups and nominal interest rates in the data. However, it is worth noting that the performance of the method does not depend significantly on the relative dispersion of structural shocks.

Table 1: Parameter Values of Baseline Model

Description	Parameter	Value
Discount factor	β	0.99
Inverse Frisch Elasticity	φ	1
Elasticity subtitution labor	ϵ_w	7
Habit formation	h	0.6
Output labor elasticity	$1-\alpha$	2/3
Calvo parameter	θ_w	0.75
Taylor rule inflation	ϕ_π	1.5
Long-run GDP growth	g_y	0.02
SD of TFP shock	σ_a	0.817
SD of monetary shock	σ_ν	0.24
SD of discount factor shock	σ_z	3.25

Note: Parameters used in model's Monte Carlo Simulations. The values are generally standard in the literature. The standard deviation of TFP shocks matches that of Fernald (2012) TFP series. The SD of the discount factor and monetary shock are set to match the SD of the wage markup and nominal interest rate in the data.

Using the calibrated model, we perform 10,000 simulations of a length of 70 years and get simulated series of GDP growth and wage markups. We then estimate SVAR in equation (10) using TFP shocks as an IV to estimate matrix $\mathbf{C}$, and get as a result estimated potential GDP growth rates. We assess the precision of our method by comparing our estimated series with the true ones coming from the model.

The results are shown in table 2, which displays the median value of key statistics for the distribution of potential GDP growth. The method is able to estimate potential GDP growth with high precision. The distributions of true and estimated potential GDP growth are very similar in

Table 2: Testing the Method

	Model Δy_t^p	Estimated Δy_t^p
Standard Deviation	0.578	0.575
Minimum	-1.620	-1.614
Maximum	1.622	1.614
Correlation with Model Δy_t^p		0.997

Note: The numbers are median values across the 10,000 simulations.

the sense that there are small differences in (median) standard deviation, minimum and maximum values. Moreover, the two series have a median correlation of 0.997 across all simulations. This implies that it is possible to precisely infer potential output consistent with a set of structural models with a simple SVAR estimation.

4 Estimating US Potential GDP

In this section we apply the baseline method to US data. We first describe the data series used in the estimation and then detail estimation results using different identification alternatives. Finally, we explain the shock identification strategy implicit in our method and discuss policy and business-cycle implications of our results.

4.1 Data

To estimate the SVAR suggested by proposition 1 we use series of GDP growth and wage markup for the period 1950Q1-2019Q4. In order to get a wage markup series, we assume a log utility (a preference consistent with a BGP) and a Firsch elasticity equal to 1. We use the following formula,

$$\mu_t^{w,data} = \log\left(\frac{W_t^{data}}{P_t^{data}}\right) - \log\left(N_t^{data}\right) - \log\left(C_t^{data}\right)$$

where W_t^{data}/P_t^{data} is real wage data, N_t^{data} is hours per capita and C_t^{data} is a series of total consumption per capita.[11] Computed in this way, the wage markup series has a clear low-frequency component with a downward trend, possibly reflecting low-frequency changes in preferences, or even

[11] Appendix D describes the data series employed in detail.

demographic or institutional changes affecting the labor market. Following Galí et al. (2007), we detrend the series using a third-order polynomial, because our focus is on business cycle fluctuations of potential GDP.

We use three different instruments or proxies for estimating SVAR in equation (10). First, we use Fernald (2012) series that measures the quarterly growth rate in total factor productivity as an IV for TFP shocks in our model ε_{at}. Second, we use for robustness Romer and Romer (2004) monetary shocks as an instrument for the linear combination of demand shocks ξ_t. Third, we use an alternative measure of TFP from Comin et al. (2020) that employs a new method removing various simplifying assumptions in Fernald (2012). In the rest of the paper we use Fernald (2012) TFP proxy as our baseline choice and employ the other two instruments for robustness checks. We prefer Fernald (2012) proxy over the monetary shock series because the latter tends to be a weaker instrument.[12] Moreover, Comin et al. (2020) TFP series is a relevant instrument but it has the shortcoming that its sample is short (it starts in 1996) and it is at annual frequency.

Lastly, to reduce concerns of possible bias due to spurious low-frequency correlation between our series, we detrend real GDP and Fernald's TFP series using a third-order polynomial. Because of its short sample, we used a linear trend for Comin et al. (2020) TFP series.[13]

4.2 SVAR-IV

Using the data detailed in the previous section we estimated the VAR implicitly defined by system (10) using OLS equation by equation.[14] This estimation provided a consistent estimate of matrix **B**. Then, as a second step, we inferred matrix **C** using the typical constraints in a SVAR-IV. We reproduce this identifying constraints below for completeness.[15] First, letting $u_t^{\Delta y}$ and $u_t^{\mu^w}$ be the

[12] To minimize endogeneity concerns, following Romer and Romer (2004), we apply the "recursiveness" assumption and use monetary shocks lagged one quarter as instruments in the SVAR-IV.

[13] See Fernald (2007) for an analysis of how low frequency correlation can affect SVAR results. The results barely change if we use a linear or quadratic trends instead of third-order polynomials.

[14] Note that there might be a bias in OLS estimates if structural shocks are serially correlated. However, we did not find evidence on serial correlation of estimated residuals using US data. If, when using data from other countries, OLS residuals happen to be serially correlated then the VAR estimation must include a filtering step to deal with the bias - see Cochrane and Orcutt (1949) and related methods.

[15] See more details in Stock (2008).

VAR residuals or forecast errors, the following equation holds,

$$\begin{bmatrix} u_t^{\Delta y} \\ u_t^{\mu^w} \end{bmatrix} = \mathbf{C} \begin{bmatrix} \varepsilon_{at} \\ \xi_t \end{bmatrix} \tag{11}$$

Letting Σ_u be the variance-covariance matrix of residuals, it is true that $\Sigma_u = \mathbf{CC}'$ which implies three different identifying restrictions. SVAR-IV uses external instruments to get the forth constraint to be able to identify all four elements of $\mathbf{C}$. In our empirical application we use instruments related to either shocks, but for simplicity assume that we have access to only one instrument correlated with ε_{at} and not with ξ_t. Particularly, assume we have access to a variable $\omega_t = \iota\varepsilon_{at} + \epsilon_t$, where ι is a parameter and ϵ_t represents an additive normally distributed measurement error. Hence, using this proxy variable allows us to introduce the following additional constraint,

$$\frac{\mathbb{E}(u_t^{\Delta y}\omega_t)}{\mathbb{E}(u_t^{\mu^w}\omega_t)} = \frac{c_{11}}{c_{21}}$$

where c_{ij} for $i, j = \{1, 2\}$ are matrix $\mathbf{C}$'s elements. This means that using the ratio of covariances between the residuals and the IV for ε_{at} we can infer the ratio of $\mathbf{C}$'s first column components. It is straightforward to show that in the case of a proxy for ξ_t, it is possible to infer the ratio of $\mathbf{C}$'s second column components.

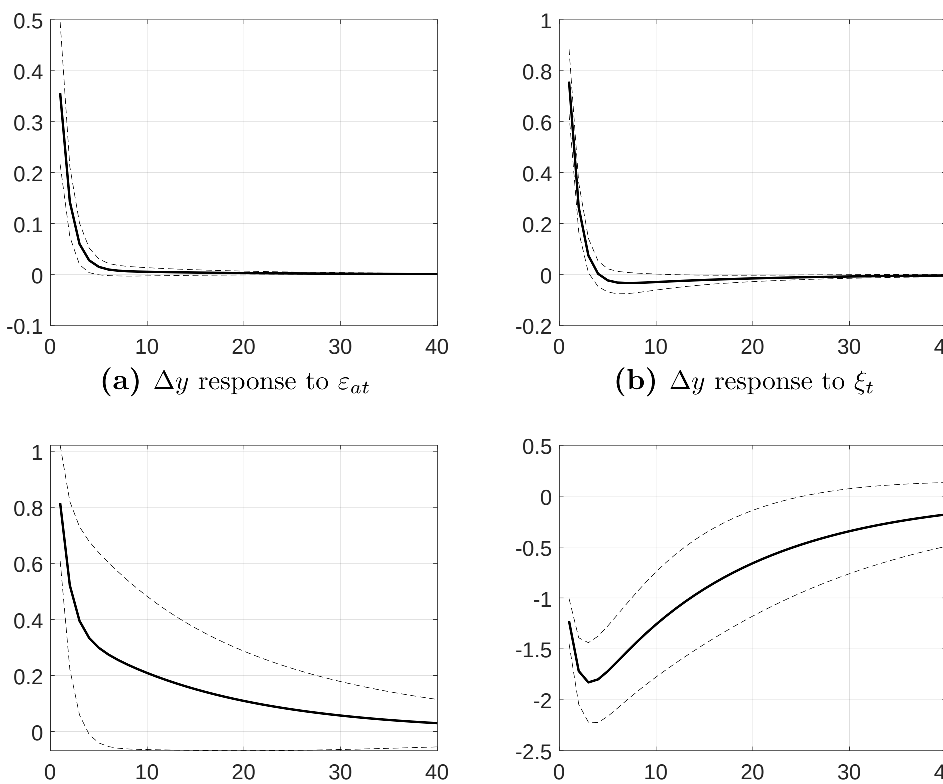
Figure 1 describes the estimation results through impulse response functions (IRFs) to one standard deviation shocks. We used Fernald (2012) TFP series as an instrument for ε_{at} in these results.[16] Figures in panel (a) and (b) show the impact on GDP growth of productivity shocks ε_{at} and the linear combination of demand shocks ξ_t. The results indicate that both shocks generate a short-run increase in GDP growth with a subsequent decline. In both cases, the impact of both shocks lasts for about 15 quarters. Figures in panels (c) and (d) show the impact of demand and productivity shocks to the wage markup. The figure in panel (d) shows a decrease in the wage markup as a consequence of demand shocks, which reflects a "tighter" labor market. On the contrary, in panel (c) shows that wage markups actually increase as a consequence of productivity shocks, implying a reduction in total hours worked. The fact that wage markups increase with

[16] Results are qualitatively similar if we use Romer and Romer (2004) monetary policy shocks or Comin et al. (2020) TFP.

productivity shocks is consistent with a set of contributions documenting a decrease in hours or an increase in unemployment with positive productivity shocks, such as, Blanchard and Quah (1989), Galí (1999), Francis and Ramey (2005), Basu et al. (2006) and Fernald (2007).

Notice that we are estimating the SVAR and plotting the IRFs in figure 1 under the "null hypothesis" that productivity shocks are not correlated with contemporaneous or lagged demand shocks. We analyze the plausibility of that lack of correlation in section 6 when we test whether potential GDP, a variable that depends on ε_{at}, reacts to different demand shocks.

Figure 1: IRFs to one standard deviation shocks



(a) Δy response to ε_{at} **(b)** Δy response to ξ_t

(c) μ^w response to ε_{at} **(d)** μ^w response to ξ_t

Note: IRFs to one standard deviation shocks (response is in percentage points; time variable is in quarters). The dashed lines represent bootstrap two standard error confidence intervals. IRFs come from a proxy-SVAR using Fernald (2012) TFP series as instrument. Sample: 1950Q1-2019Q4.

4.3 Results

Figure 2 depicts our output gap series (log difference between GDP and potential GDP) computed using the whole sample (1950Q1-2019Q4) and Fernald (2012) TFP series as a proxy in the SVAR estimation. The figure also shows CBO's output gap series, a widely used series in economic analysis.[17] In this figure, and throughout the rest of the paper, we follow Coibion et al. (2017b) and assume that potential GDP was equal to actual GDP in 2006Q3, a quarter for which CBO estimated a null gap. Note that we need this assumption to determine the level of potential GDP from the estimated growth rates.

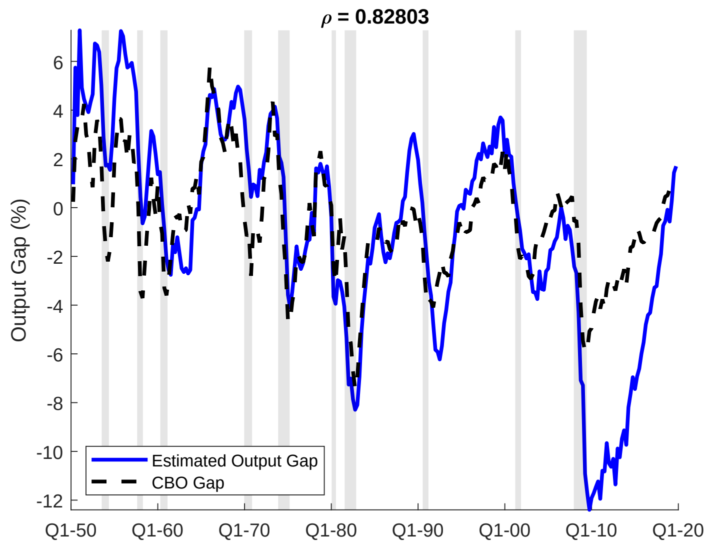
Interestingly, figure 2 shows that the two series are highly correlated with a correlation coefficient of 0.83. Nevertheless, there are stark differences around some recessions; most notably, during and after the Great Recession. Specifically, around 2009, our output gap series points to a gap lower than -10%, whereas the CBO calculates a gap of around -5%. Further, our series indicates a much faster increase in the gap after the recession compared to CBO's counterpart.

Why is there such a big difference between the two output gap series during and after the Great Recession? Figure 3 shows the evolution of potential GDP predicted by our method. The figure shows results using as a proxy Fernald (2012) TFP series in panel (a), the ones using Romer and Romer (2004) monetary policy shocks in panel (b), and series estimated using Comin et al. (2020) TFP in panel (c). All panels compare potential GDP with CBO's estimates and actual GDP during and after the Great Recession. We show in light blue lines various estimates using samples with different starting points, from 1950Q1 to 1990Q1. The thicker dark blue line is the median across all estimates.

As it is clear from our results, there is an important difference in the evolution of potential GDP estimates between ours and CBO's. In particular, our results imply a jump in potential GDP during the recession and a subsequent disappointing growth afterwards. The sharp increase in potential GDP explains the large drop in the output gap in 2009, and the low potential growth rate afterwards is related to the rapid increase in the gap from 2010. This evolution of potential GDP is directly related to the path of TFP during and after the recession. In fact, both Fernald (2012) and

[17] The CBO computes potential GDP using a non-structural method. In particular, they assume a Cobb-Douglas production function and extract the "business cycle" component of each input (including the Solow residual) using reduced form regressions. See Shackleton (2018) for details.

Figure 2: Output Gaps: Baseline Method vs CBO



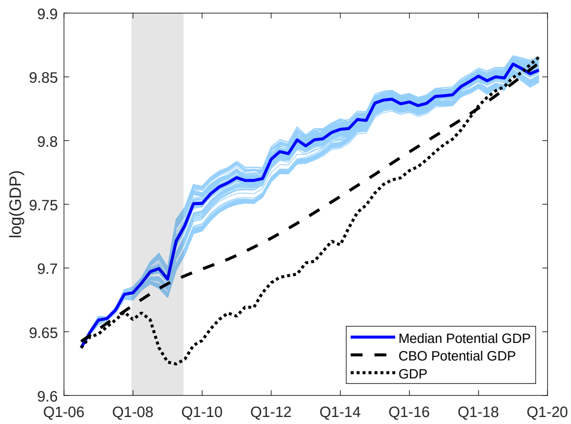
Note: Output gap is defined as the log difference between GDP and potential GDP. Fernald (2012) TFP series are used as a proxy to estimate potential output. The value of ρ indicates the correlation between the two series. See appendix D for data details.

Comin et al. (2020) TFP series indicate a sudden increase in productivity during the recession and low growth rates afterwards. Moreover, as figure 3 highlights, our prediction is robust to different estimation strategies and samples used. The results using either Fernald (2012), Romer and Romer (2004) or Comin et al. (2020) proxies are qualitatively similar and do not seem to be affected by changes in sample lengths. Interestingly, both Romer and Romer (2004) and Comin et al. (2020) proxies imply a higher increase in TFP during the great recession but a significantly lower average growth rate afterwards, compared to the results using Fernald (2012) TFP.

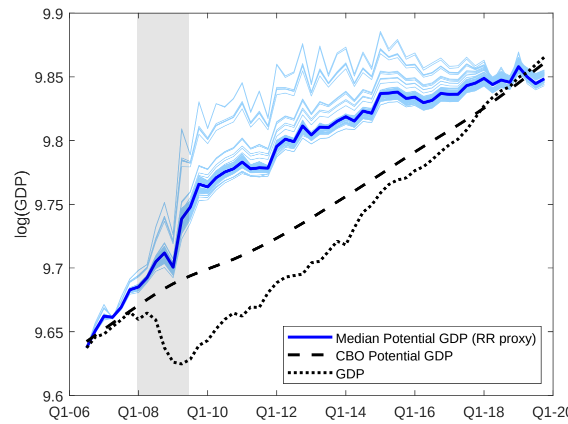
The story that CBO's estimates tell is quite different. They point to a gradual slowdown in potential GDP growth rate during and after the recession, but they also estimate a higher and increasing potential growth rate after 2012. This explains why CBO's output gap increases at a much slower rate in figure 2.[18]

[18]In the online appendix, we perform real-time estimations using vintage data and compare our estimates with CBO's. The real-time results also display important differences during and after the Great Recession. Our estimates point to larger output gaps as data is revised and new data becomes available, whereas CBO series suggests smaller output gaps.

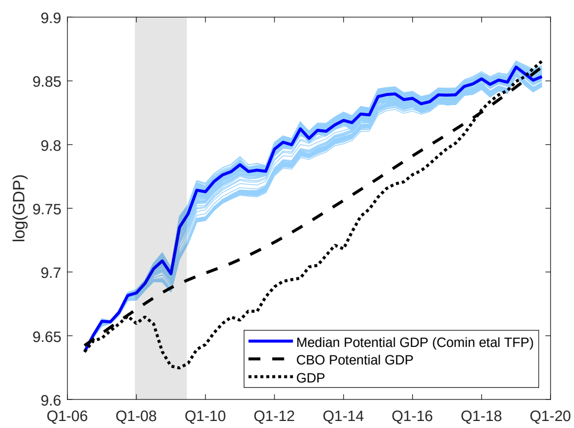
Figure 3: Potential Output During and After the Great Recession



(a) Fernald (2012) TFP as proxy



(b) Romer and Romer (2004) shocks as proxy



(c) Comin et al. (2020) shocks as proxy

Note: The figures show potential GDP series, compared with CBO's series and actual GDP. Panel (a) shows estimates using Fernald (2012) TFP as a proxy, panel (b) uses Romer and Romer (2004) monetary shocks, and panel (c) uses Comin et al. (2020) TFP. Light blue lines highlight estimates using different samples starting at different dates, using as starting dates quarters from 1950Q1 to 1990Q1. The thick dark blue line represents the median value across all estimates. For the results using Romer and Romer (2004) monetary shocks we excluded samples starting after 1974Q1 because monetary shocks are not relevant instruments in those cases. See data appendix D for details.

4.4 On Structural Shocks Identification

Why do our results predict an increase in TFP and potential GDP during the Great Recession? As we mentioned before, the increase in productivity during 2008 and 2009 is consistent with the increase in Fernald (2012) and Comin et al. (2020) measures of TFP. However, in our SVAR-IV we assumed that our instruments were subject to measurement error and possibly biased. Hence, in principle, the increase in the IVs for TFP could be regarded as error by our method. Consequently, to answer this question we need to clarify how productivity and demand shocks are identified in our SVAR. To get intuition we use system (11) to infer the following relationship between forecast errors or VAR residuals $u_t^{\Delta y}$ and $u_t^{\mu^w}$,

$$u_t^{\Delta y} = \frac{c_{12}}{c_{22}} u_t^{\mu^w} + \theta_0 \varepsilon_{at} \tag{12}$$

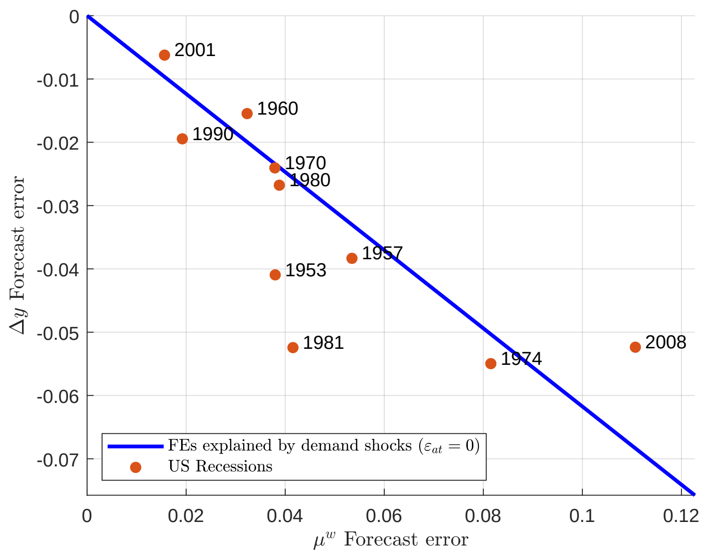
where θ_0 is defined in (9) and c_{ij} for $i, j = \{1, 2\}$ are matrix $\mathbf{C}$ elements. Each element of $\mathbf{C}$ represents the impact of every shock to either output growth or the wage markup. Therefore, by looking at figure 1 we can infer that $\frac{c_{12}}{c_{22}} < 0$ and $\theta_0 > 0$ because the impact of demand shocks on μ_t^w is negative ($c_{22} < 0$) and the rest of the IRFs show positive impacts implying positive c_{11}, c_{12} and c_{21}.

Assuming no productivity shocks, or $\varepsilon_{at} = 0$, equation (12) implies a negative relationship between growth and wage markup forecast errors. This relationship is reminiscent of the Okun (1962) law, but instead of relating the levels of output and unemployment, it represents a relationship between GDP growth and wage markup forecast errors. Equation (12) states that, in the absence of productivity shocks, growth and wage markup forecast errors should be aligned on a line passing through the origin with a slope equal to $\frac{c_{12}}{c_{22}} < 0$. Importantly, the equation reveals how the SVAR identifies productivity shocks. Specifically, any deviation from the line assuming $\varepsilon_{at} = 0$ is explained by positive or negative ε_{at} values. For instance, if forecast errors are such that $u_t^{\Delta y} > \frac{c_{12}}{c_{22}} u_t^{\mu^w}$ in a particular quarter, then our method infers that there was a positive productivity shock.

We present the estimated relationship between forecast errors in figure 4, where we used the whole sample 1950Q1-2019Q4 and matrix $\mathbf{C}$ was estimated using Fernald's TFP proxy. The blue line shows the relationship between forecast errors under the assumption of no productivity shocks. The red dots highlight the sum of forecast errors in each US recession. As it is clear from the figure, most economic recessions are either on or below the blue line implying that recessions are generally characterized either by demand shocks only or by both negative productivity and demand shocks.

However, one clear outlier in the figure is the Great Recession that is significantly above the blue line implying an important increase in productivity and potential GDP during these quarters.

Figure 4: Identification of productivity shocks



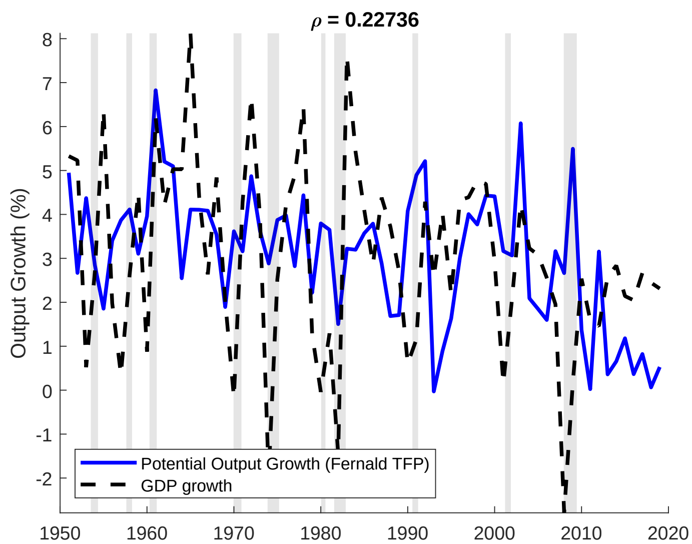
Note: relationship between forecast errors estimated through a proxy-SVAR using as proxy Fernald's TFP. The dots highlight the sum of forecast error during each US recession since 1950.

4.5 Policy and Business-Cycle Implications

Our estimates of flexible-price output have implications in terms of monetary policy. In particular, absent high-frequency shocks to markups, monetary theory suggests that the monetary authority should aim to close the output gap. Hence, a natural question to ask is how much output volatility would be reduced if the Fed carried out this optimal policy rule. We present figure 5 to answer that question; it shows series of actual and potential GDP annual growth, where potential growth is clearly less volatile than the actual one. In particular, our results indicate a potential growth standard deviation of 1.5%, whereas the volatility of actual GDP growth is 2.1%. Hence, optimal monetary policy can in principle reduce output volatility by 30%, omitting issues related to the zero lower bound that might make optimal policy unimplementable. As a consequence, optimal monetary policy is far from eliminating business cycle fluctuations.

Finally, figure 5 highlights a fairly low correlation between actual and potential GDP with a correlation coefficient of 0.23. This result provides evidence on the importance of labor wedge fluctuations shaping the business cycle, in line with Chari et al. (2007) and Galí et al. (2007).

Figure 5: Actual and Potential GDP Annual Growth Rates



Note: We used Fernald (2012) TFP series as a proxy to estimate potential output growth rate. The value of ρ indicates the correlation between the two series. See data appendix D for details.

5 Extensions

This section analyzes basic extensions to our baseline method. We first show how our approach changes when we add price rigidities to the underlying model. Second, we address the importance, in terms of method performance, of the omission of endogenous capital stock in the baseline method.[19]

[19]We worked other extensions such as adding government spending shocks or a COVID-19 labor supply shock. We refer the reader to the online appendix for details and results.

5.1 Price Rigidities

There is growing consensus in the literature indicating that fluctuations in wage markups explain most of the labor wedge variance; some contributions in this area are Galí et al. (2007), Hall (2009), Shimer (2009), Karabarbounis (2014), and Nekarda and Ramey (2020).[20] This sequence of contributions is the main reason why we excluded price rigidities from the baseline model. However, we can show that it is straightforward to change our method and incorporate a cyclical price markup.

Introducing price rigidities implies changing equation (5) in the baseline model for the following one,

$$w_t - \frac{1}{1-\alpha} a_t + \frac{\alpha}{1-\alpha} y_t = -\mu_t^p \tag{13}$$

where μ_t^p represents the price markup in the economy. Having nominal rigidities in goods markets also adds a price Phillips curve,

$$\pi_t = -\kappa \mu_t^p + \beta \mathbb{E}_t \pi_{t+1} \tag{14}$$

where κ represents the slope parameter. The rest of the model remains unchanged, and potential GDP growth is defined as in (9). Defining the (log) labor wedge as $\tau_t \equiv \mu_t^w + \mu_t^p$, proposition 3 describes the method to estimate potential GDP under wage and price rigidities.

Proposition 3. *Using a model with price and wage rigidities, θ_0, θ_1 and ε_{at} in equation (9) can be estimated from the following the estimation of the system,*

$$\begin{bmatrix} \Delta y_t \\ \tau_t \end{bmatrix} = \mathbf{B} \begin{bmatrix} \Delta y_{t-1} \\ \mu_{t-1}^w \\ \mu_{t-1}^p \end{bmatrix} + \mathbf{C} \begin{bmatrix} \varepsilon_{at} \\ \xi_t \end{bmatrix} \tag{15}$$

where ξ_t is a weighted average of demand shocks. In particular, letting c_{ij} and b_{ij} for $i, j = \{1, 2\}$

[20]Recently, Bils et al. (2018) challenged this view claiming that price markups are at least as cyclical as wage markups.

be the elements of the first two rows and columns of matrices $\mathbf{B}$ *and* $\mathbf{C}$*,*

$$\theta_0 = c_{11} - \frac{c_{21}c_{12}}{c_{22}} \qquad \theta_1 = b_{11} - \frac{b_{21}c_{12}}{c_{22}}$$

And ε_{at} *can be calculated using forecast errors and* $\mathbf{C}$*.*

Proof. See Appendix C. □

As the proposition describes, there is only a small modification in the system of two equations to be estimated. Now we need series of price and wage markups (μ_t^p and μ_t^w) to perform the estimation. Nevertheless, the strategy to estimate potential output from the estimated matrices is unchanged.

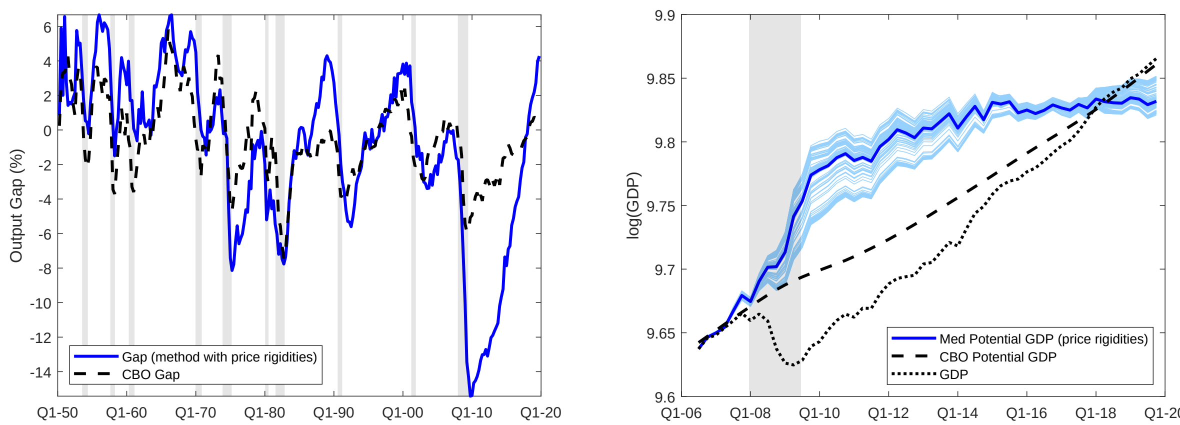
In order to apply this method, we measure price markups in the data assuming that the production function is Cobb-Douglas and, therefore, price markups are defined as the log-difference between labor productivity (output over hours) and the real wage. That is,

$$\mu_t^{p,data} = \log\left(Y_t^{data}\right) - \log\left(N_t^{data}\right) - \log\left(\frac{W_t^{data}}{P_t^{data}}\right)$$

where Y_t^{data} is GDP per capita, N_t^{data} represents total hours per capita, and $\frac{W_t^{data}}{P_t^{data}}$ is a series of real wage. As we do for wage markups, we detrend price markup series with a third-order polynomial to omit low-frequency movements that are not explained by the business cycle.

The estimated potential GDP series using this modified method are shown in figures 6. Our previous results are confirmed. The estimated output gap is highly correlated with the one computed by the CBO, with a correlation coefficient of 0.78. We still find a large difference between the two series during and after the Great Recession. As panel (b) in the figure shows, there is a different evolution of potential GDP in this period compared to CBO's predictions. Like in our baseline case, the results indicate an increase in potential GDP during the recession and a poor growth rate afterwards.

Figure 6: Including Price Rigidities: Output Gap and Potential GDP



(a) Output Gap vs CBO

(b) Potential GDP during Great Recession

Note: Panel (a) shows the output gap computed using the method assuming wage and price rigidities in the underlying model. We also include CBO's output gap for comparison. Panel (b) shows potential GDP series computed with the same method. Light blue lines in panel (b) highlight different estimates using different samples starting at different dates, using as starting dates quarters from 1950Q1 to 1990Q1. The thick dark blue line represents the median value across all estimates. See data appendix D for details.

5.2 Endogenous Capital

We showed in proposition 2 that the baseline method is consistent with a production function with capital such as $Y_t = A_t(U_t\bar{K}_t)^{\alpha}N_t^{1-\alpha}$, where the capital stock $\bar{K}_t$ is assumed to be exogenously determined. This is clearly better than assuming no capital in production, but it is still unrealistic. The main reason behind our simplifying assumption is that incorporating an endogenous capital stock into the model complicates the method in a significant way. Most importantly, this modification adds state variables that are difficult to measure and need to be incorporated as observables in our method. The new state variables to be added depend on specific assumptions, but notable examples are the capital stock of the economy and, if there are capital adjustment costs, the marginal Tobin's Q.

Hence, instead of complicating the method, in this section we test whether omitting the fact that the capital stock is endogenous represents a significant problem for the accuracy of our method. We do so by using the model proposed by Smets and Wouters (2007) (henceforth, SW) as a laboratory.

The exercise we perform is the following. We simulate data from the estimated model suggested by SW and then use it to calculate potential GDP using our method with both price and wage rigidities. We then compare the distributions of the estimated potential GDP growth and the one coming from SW model. The results of this exercise are shown in table 3, which reports the median statistics after running 10,000 simulations of a length of 70 years.

Table 3: Testing the Method with Smets and Wouters (2007) model

	Model Δy_t^p	Estimated Δy_t^p
(A) Original SW Model		
Standard Deviation	0.751	0.788
Minimum	-2.029	-2.126
Maximum	2.033	2.127
Correlation with SW Δy_t^p		0.92
(B) SW without short-run wealth effects		
Standard Deviation	1.155	1.079
Minimum	-3.117	-2.920
Maximum	3.123	2.929
Correlation with SW Δy_t^p		0.99

Note: The numbers are median values across the 10,000 simulations. Panel A uses as data generating process Smets and Wouters (2007) estimated model. Panel B uses a modified version of that model without wealth effects.

As table 3 shows in panel (A), our method's estimates are highly correlated with the true potential GDP growth coming from the model (with a correlation coefficient of 0.92). Moreover, the distribution of estimated potential growth rates seems to be close to the distribution of model simulations.

The correlation between model and estimated series in panel A is high but significantly lower than one, which points to a bias in our estimates. What are the main factors generating this bias? Omitting the endogeneity of the capital stock generates two potential sources. First, by assuming a constant detrended level of capital, our method is underestimating the variance of the capital stock which in turn reduces the variance of estimated potential output relative to the true one. Second, because investment is endogenous in a model with endogenous capital, consumption and output are not proportional anymore. This implies that the consumption-output proportionality assumed in our baseline method does not hold in Smets and Wouters (2007). This difference in the behavior of consumption in equilibrium is the second source of error. In particular, it introduces a bias in the estimation of potential hours worked, because consumption determines the short-run impact of wealth effects on labor supply.

Hence, omitting the endogeneity of capital introduces a bias in both the capital stock and potential hours. However, because the capital stock does not fluctuate significantly at a business-cycle frequency, it turns out that the most important source of bias has to do with errors measuring short-run wealth effects in the labor market. In other words, the main problem is related with the estimation of potential hours worked. This implies that the bias of our method should decrease when short-run wealth effects are less important. In particular, in a case where wealth effects are null, the estimates of potential hours should not be significantly biased. We provide evidence supporting this conclusion in the following way. First, we modify the original SW model to eliminate wealth effects. Specifically, we introduce a tax/subsidy Θ_t on household's labor income, financed through lump-sum taxes, and designed to eliminate short-run shifts in the labor supply. Consider the labor supply decision of the household in SW described by the following first order condition,

$$\left(C_t - hC_{t-1}\right) N_t^{\varphi} = \frac{W_t^h}{P_t} \tag{16}$$

where W_t^h is the nominal wage labor unions pay to households and, as before, C_t, N_t and P_t represent consumption, total hours, and the price level. To get rid of short-run wealth effects we assume a tax/subsidy on wage income Θ_t such that,

$$\Theta_t = \frac{\left(C_t - hC_{t-1}\right)}{\left(\bar{C}_t - h\bar{C}_{t-1}\right)}$$

where $\bar{C}_t$ is consumption at the BGP. Therefore, including the tax, the labor supply becomes

$$\left(C_t - hC_{t-1}\right) N_t^{\varphi} = \Theta_t \frac{W_t^h}{P_t}$$

Combining the last two equations, it is clear to see that short-run wealth effects disappear,

$$\left(\bar{C}_t - h\bar{C}_{t-1}\right) N_t^{\varphi} = \frac{W_t^h}{P_t}$$

We carry out the simulation exercise again using this modified version of SW with the originally estimated parameters. The results of this second exercise are shown in panel (B) in table 3. The table shows an almost perfect correlation between the original series and the estimates using our method. This clearly indicates that the main source of bias behind our assumption of capital stock exogeneity has to do with the presence of short-run wealth effects. As a consequence, if one believes that these effects are not meaningful, then one should expect a small bias from assuming exogenous capital accumulation.

6 Is Potential GDP affected by demand shocks?

Traditional models of business-cycle fluctuations that started with the Real Business Cycle Theory assume that productivity is an exogenous process that is not affected by current or past demand shocks. Productivity is typically the only source of low-frequency fluctuations in these models. This implies that, according to this approach, the behavior of GDP in the medium or long-run is not affected by demand shocks such as fiscal or monetary policy shocks.

An alternative view suggests that the behavior of GDP in the medium or long run can be affected by cyclical fluctuations. According to this interpretation, past or current demand shocks can generate persistent effects on GDP, a result called output hysteresis in the literature.

There are several theories that point to different mechanisms generating output hysteresis. One relevant set suggests that GDP persistence happens because potential GDP is affected by demand shocks. These theories emphasize that demand shocks can affect total factor productivity through learning by doing, human capital and knowledge accumulation, or the speed of adoption of new technologies.[21] In this section we test these theories using our estimated series of potential output. Specifically, we estimate impulse response functions (IRFs) of potential GDP to demand shocks. Traditional models of business cycles predict that there should not be an impact of demand shocks on potential GDP, whereas TFP-related hysteresis hypotheses imply the opposite.

We employ local projection methods introduced by Jordà (2005) to compute the IRFs. In particular, letting ε_{dt} be a demand shock, the response of potential GDP (Y^p) at horizon h is

[21] See, for example, Stadler (1990), Stiglitz (1993), Fatás (2000), and Comin and Gertler (2006).

estimated by running the regression:

$$\log\left(Y^p_{t+h}\right) = \theta_h \varepsilon_{dt} + \text{Control Variables} + e_{t+h} \tag{17}$$

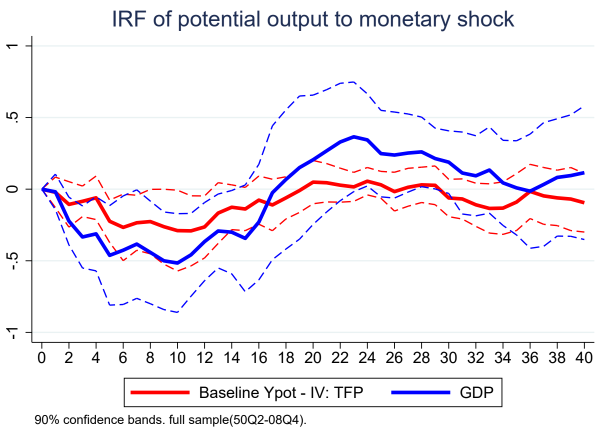
where e_{t+h} is an error term, Control Variables include a cubic trend, dummy variables controlling for the Great Recession and 9/11, and four lags of Y^p in logs and ε_{dt}. The estimate of θ_h measures the response of potential GDP at horizon h in percentage points.

We analyze four different demand shocks: Romer and Romer (2004) monetary policy shock, fiscal policy shock estimated by Auerbach and Gorodnichenko (2012), Ramey (2016) military spending shock, and our SVAR demand shock ξ_t. The first three shocks are known for generating persistent effects on GDP, even though they themselves are not persistent. The estimated IRFs showing the impact of one standard deviation shocks are shown in figure 7.

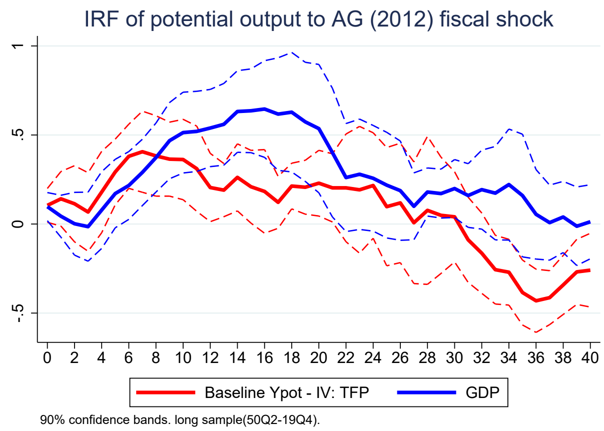
The figure clearly shows a statistically significant impact of demand shocks on potential GDP. For the cases of monetary shock and fiscal shock in panel (a) and (b), potential output seems to be affected around five quarters after the shock hits the economy. Actual GDP moves in the same direction as potential GDP, which implies that business cycle fluctuations generated by these demand shocks increase potential output during booms and decrease it in recessions. The IRFs provide evidence of a persistent impact on potential GDP. Monetary shocks affect potential output for 4 years, whereas fiscal shocks generate an increase for 5 years approximately. Interestingly, our results imply that the effect of demand shocks on potential GDP is not permanent. In other words, negative (positive) demand shocks generate a long period of low (high) potential GDP growth that is followed by a reversal period with higher (lower) growth. This result seems to be at odds with theories that suggest a permanent effect on potential GDP, such as, theories related to R&D investment. On the other hand, it is consistent with theories suggesting temporary effects like the ones related to technology diffusion, or learning by doing. However, it should be noted that this lack of persistence might be related with the average size of shocks used in this exercise. Indeed, evidence seems to suggest that permanent effects on potential GDP are related with deep recessions.[22] Hence, it might be the case that our results show only temporary effects on potential GDP because of the linearity of our estimated model.

[22] See Blanchard et al. (2015) and Anzoategui et al. (2019).

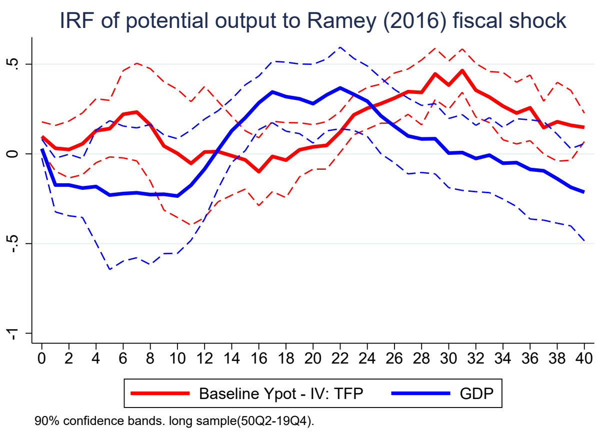
Figure 7: Potential Output Response to Demand Shocks



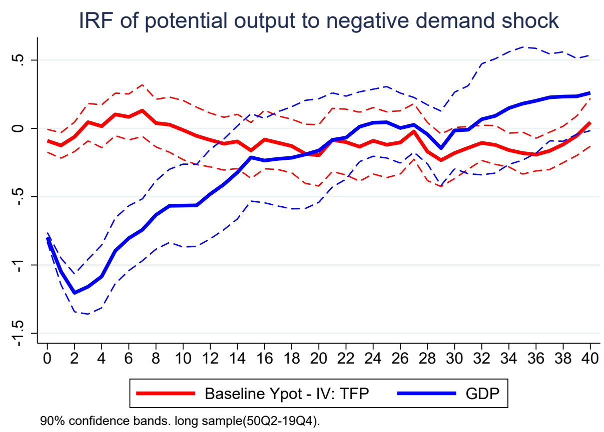
(a) Romer and Romer (2004) monetary shock



(b) Auerbach and Gorodnichenko (2012) fiscal shock



(c) Ramey (2016) defense spending shock



(d) SVAR demand shock ξ_t

Note: figures show potential GDP IRFs to one standard deviation demand shocks: Panel (a) Romer and Romer (2004) monetary policy shocks, panel (b) fiscal policy shocks from Auerbach and Gorodnichenko (2012), and panel (c) defense spending shocks from Ramey (2016). 90% confidence intervals are reported. Potential output was computed by our baseline method using as proxy Fernald (2012) TFP series and the sample from 1950Q1 to 2019Q4.

Panels (c) and (d) in figure 7 show the effect of a positive defense spending shock and a negative ξ_t shock, respectively. Both panels show evidence of an effect of demand shocks on potential GDP. However, differently from panels (a) and (b), the shocks have a large delay in its effect on potential output. In particular, these IRFs point to a delay of at least 14 quarters approximately, whereas panels (a) and (b) imply a 5 quarters lag.

Overall, the results support the theories claiming that potential output is endogenous and responds to business cycle fluctuations increasing the persistence of GDP. Moreover, how fast demand shocks are able to affect potential GDP meaningfully seems to depend on the type of the shocks being considered.

TFP-related hysteresis hypotheses claim that TFP is endogenous. We provide further evidence supporting these hypotheses analyzing the effect of demand shocks on TFP. Figure 8 show the response of Fernald (2012) TFP series to the four demand shocks mentioned above. The IRFs were computed using equation (17) with (log) TFP being the left hand side variable instead of potential output. The four different figures clearly show an endogenous response of TFP. Contractionary monetary policy shocks tend to reduce TFP with a delay of 5 quarters, whereas expansionary fiscal shocks and defense spending shocks increase TFP with a delay of one year and 4 years, respectively. Our SVAR demand shock, ξ_t, generates a persistent decline in productivity 3 years after the shock hits. The results again suggest a temporary impact on TFP, which is consistent with our previous finding that demand shocks have a persistent but temporary effect on potential output.

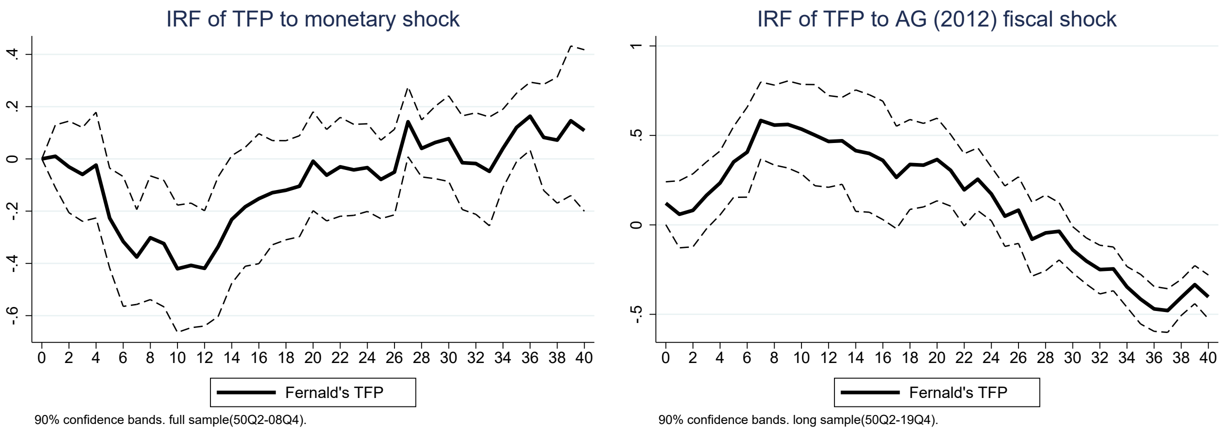
7 Methodological issues related to endogenous TFP

Section 6 provided evidence indicating that demand shocks affect productivity growth with a delay. However, our baseline method assumed productivity was exogenous and uncorrelated with any other shocks. In this section, we relax that assumption and modify our method to be consistent with the findings in section 6. In particular, we remain agnostic on the specific channel through which demand shocks can affect productivity and assume that productivity growth, ε_{at}, is determined by the following expression,

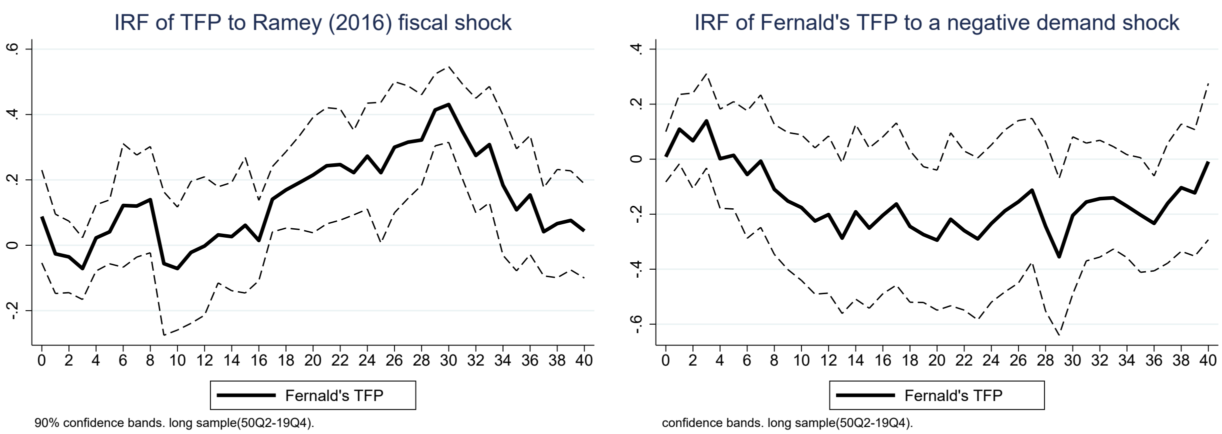
$$\log(\tilde{A}_t) - \log(\tilde{A}_{t-1}) = \varepsilon_{at} = \eta_a \tilde{\varepsilon}_{at} + \sum_{j=n_1}^{n_2} \eta_j \xi_{t-j} \tag{18}$$

where $\tilde{A}_t$ is detrended TFP level, $\tilde{\varepsilon}_{at}$ is a truly exogenous TFP shock, and the η_j are parameters setting weights on past demand shocks ξ_{t-j}. We assume $\sum_{j=n_1}^{n_2} \eta_j^2 + \eta_a^2 = 1$ to keep the variance of ε_{at} equal to one. Expression (18) states that productivity growth can be affected by truly exogenous

Figure 8: TFP Response to Demand Shocks



(a) Romer and Romer (2004) monetary shock **(b)** Auerbach and Gorodnichenko (2012) fiscal shock



(c) Ramey (2016) defense spending shock **(d)** SVAR ξ_t shock

Note: figures show TFP IRFs to one standard deviation demand shocks: Panel (a) Romer and Romer (2004) monetary policy shocks, panel (b) fiscal policy shocks from Auerbach and Gorodnichenko (2012), and panel (c) defense spending shocks from Ramey (2016). 90% confidence intervals are reported. We use Fernald (2012) TFP series and a sample from 1950Q1 to 2019Q4.

TFP shocks ($\tilde{\varepsilon}_{at}$) and demand shocks (ξ_t). Specifically, ξ_t is assumed to affect productivity for $n_2 - n_1$ quarters with a lag of n_1 quarters.

We opted for this simple way of incorporating endogenous productivity because of two reasons. First, this specification is directly related to our findings in section 6 and, therefore, we are able to set values for n_1 and n_2 based on those results. Second, introducing equation (18) to our baseline model does not imply a change in the method to compute potential GDP. With this simple modification, the strategy is still the one described in proposition 1: a SVAR estimation. Specifically, the SVAR to be estimated is now,

$$\begin{bmatrix} \Delta y_t \\ \mu_t^w \end{bmatrix} = \mathbf{B} \begin{bmatrix} \Delta y_{t-1} \\ \mu_{t-1}^w \end{bmatrix} + \mathbf{C} \begin{bmatrix} \eta_a \tilde{\varepsilon}_{at} + \sum_{j=n_1}^{n_2-1} \eta_j \xi_{t-j} \\ \xi_t \end{bmatrix} \tag{19}$$

The only problem related to the incorporation of endogenous productivity has to do with the estimation of matrix $\mathbf{B}$ in system (19). As can be noted in the system, ε_{at} now may be correlated with Δy_{t-1} and μ_{t-1}^w implying that OLS estimates are biased. To assess the magnitude of the problem we performed Monte Carlo simulations using as data generating process our baseline estimates. We set the delay in the response of ε_{at} to demand shocks equal to the shortest one we estimated in section 6, that is, 5 quarters (implying $n_1 = 5$). We also assumed that $n_2 = 6$ for simplicity and set $\eta_5 = 0.005$ to match a change of 0.5% in Fernald's TFP after a demand shock, which is the largest change implied by figures 8. Our results, detailed in our online appendix, indicate the presence of a bias that seems quantitatively small. This relatively unimportant bias is due to the size of η_5 and the delay between the demand shock and its effect on TFP, which reduces the correlation between ε_{at} and Δy_{t-1} or μ_{t-1}^w. In fact, the longer the delay (higher n_1), the less important the bias should be.

The simulation results point to a small bias and, therefore, typical OLS estimation does not seem problematic for the case of the US. Nevertheless, for completeness we propose in the online appendix an alternative Generalized Method of Moments (GMM) method that can be useful when the bias due to TFP endogeneity is large. The method uses estimated lagged demand shocks ξ_t as instruments. We tested the performance of the GMM method through Monte Carlo Simulations and found that even though GMM works better when the bias is large, it shows a poorer performance when the source of bias seems to be minor. For that reason, our preferred method for the case of the US is still OLS. Hence, in conclusion, the potential GDP estimates shown throughout the paper are consistent with an underlying model in which TFP is endogenously affected by demand shocks.

8 Conclusion

We presented a new method to compute potential GDP that has several advantages over existing approaches. It provides an estimate of the counterfactual flexible-price level of output and it is not subject to the Lucas Critique. Importantly, it is simple and only requires the estimation of a SVAR. The method is consistent with a large set of models with different assumptions in terms preferences, technology, expectation formation, and monetary policy rules. Moreover, it is a very good approximation of potential GDP measures coming from more complicated models such as Smets and Wouters (2007), especially if short-run labor-supply wealth effects are unimportant.

By applying our method to the US data, we provided two important insights. First, output gap estimates are highly correlated with those computed by the CBO. However, compared to CBO's estimates, our method predicts a much lower potential growth rate after the Great Recession. Second, we showed that potential GDP in the US seems to be affected by demand shocks, a result consistent with a group of theories predicting output hysteresis.

Our results are important for economic policy design in two different aspects. First, our paper provides a new method to compute potential output that can be incorporated into the set of instruments employed in policy evaluation. Second, our finding that recessions can actually reduce potential GDP is relevant information to consider when deciding how "aggressive" monetary or fiscal policy should be during downturns.

Appendices

A Proof of proposition 1

After detrending and loglinearizing, the baseline model can be summarized by the following system,

$$-\phi_\pi \pi_t - \sigma_i \nu_t = \mathbb{E}_t \Delta m u_{t+1} - \mathbb{E}_t \pi_{t+1} + \sigma_z \left(\mathbb{E}_t z_{t+1} - z_t\right) \tag{A.1}$$

$$\Delta m u_t = -\frac{1+g_y}{1+g_y-h}\Delta y_t + \frac{h}{1+g_y-h}\Delta y_{t-1} \tag{A.2}$$

$$\pi_t^w = -\kappa^w \mu_t^w + \beta \mathbb{E}_t \pi_{t+1}^w \tag{A.3}$$

$$\mu_t^w - \mu_{t-1}^w = -\left(\frac{1+g_y}{1+g-h} + \frac{\alpha+\varphi}{1-\alpha}\right)\Delta y_t + \frac{h}{1+g_y-h}\Delta y_{t-1} + \frac{1+\varphi}{1-\alpha}\sigma_a \varepsilon_{at} \tag{A.4}$$

$$\frac{1}{1-\alpha}\sigma_a \varepsilon_{at} - \frac{\alpha}{1-\alpha}\Delta y_t = \pi_t^w - \pi_t \tag{A.5}$$

The first equation in the SVAR is equation (A.4) above. The second equation comes from the fact the the wage markup is a function of the state variables and shocks in the model. Hence,

$$\mu_t^w = \gamma_a \varepsilon_{at} + \gamma_z z_t + \gamma_\nu \nu_t + \gamma_\mu \mu_{t-1}^w + \gamma_y \Delta y_{t-1}$$

where the γ's are function of deep parameters of the model. Defining $\xi_t = \frac{\gamma_z}{\sqrt{\gamma_z^2+\gamma_\nu^2}} z_t + \frac{\gamma_\nu}{\sqrt{\gamma_z^2+\gamma_\nu^2}} \nu_t$ and $\gamma_\xi = \sqrt{\gamma_z^2+\gamma_\nu^2}$ then,

$$\mu_t^w = \gamma_a \varepsilon_{at} + \gamma_\xi \xi_t + \gamma_\mu \mu_{t-1}^w + \gamma_y \Delta y_{t-1} \tag{A.6}$$

Now combining (A.4), (A.6) and after some algebra we find the following SVAR,

$$\begin{bmatrix} \Delta y_t \\ \mu_t^w \end{bmatrix} = \mathbf{B} \begin{bmatrix} \Delta y_{t-1} \\ \mu_{t-1}^w \end{bmatrix} + \mathbf{C} \begin{bmatrix} \varepsilon_{at} \\ \xi_t \end{bmatrix}$$

where,

$$\mathbf{B} \equiv \begin{bmatrix} \frac{\frac{h}{1+g_y-h}-\gamma_y}{\frac{1+g_y}{1+g_y-h}+\frac{\alpha+\varphi}{1-\alpha}} & \frac{1-\gamma_\mu}{\frac{1+g_y}{1+g_y-h}+\frac{\alpha+\varphi}{1-\alpha}} \\ \gamma_y & \gamma_\mu \end{bmatrix} \qquad \mathbf{C} \equiv \begin{bmatrix} \frac{\frac{1+\varphi}{1-\alpha}\sigma_a-\gamma_a}{\frac{1+g_y}{1+g_y-h}+\frac{\alpha+\varphi}{1-\alpha}} & \frac{-\gamma_\xi}{\frac{1+g_y}{1+g_y-h}+\frac{\alpha+\varphi}{1-\alpha}} \\ \gamma_a & \gamma_\xi \end{bmatrix}$$

Finally, using the matrices above it is easy to check that,

$$\theta_0 = c_{11} - \frac{c_{21}c_{12}}{c_{22}} \qquad \theta_1 = b_{11} - \frac{b_{21}c_{12}}{c_{22}}$$

Also, having the matrices $\mathbf{B}$ and $\mathbf{C}$ it is easy to get the structural shocks ξ_t and ε_{at},

$$\mathbf{C}^{-1}\left\{\begin{bmatrix}\Delta y_t \\ \mu_t^w\end{bmatrix} - \mathbf{B}\begin{bmatrix}\Delta y_{t-1} \\ \mu_{t-1}^w\end{bmatrix}\right\} = \begin{bmatrix}\varepsilon_{at} \\ \xi_t\end{bmatrix} \tag{A.7}$$

for all t.

B Proof of proposition 2

The basic strategy to prove propositon 2 is to show that the method to find potential GDP does not change with every suggested modification. We do that below,

1. *Changing the production function for any other production function that only uses labor as input*

   To show that this change does not affect the method, let's assume a general production function,

   $$Y_t = A_t f(N_t) - \Theta(1+g_y)^t$$

   keeping the rest of the baseline assumptions intact. After detrending and loglinearizing as in the baseline case we find that,

   $$\mu_t^w - \mu_{t-1}^w = \Omega_1\varepsilon_{at} - \Omega_2\Delta y_t + \Omega_3\Delta y_{t-1} \tag{B.1}$$

where,

$$\Omega_1 \equiv \left[1 - \left(\frac{f''(N)N}{f'(N)} - \varphi\right)\frac{f(N)}{f'(N)N}\right]\sigma_a$$
$$\Omega_2 \equiv -\left[\left(\frac{f''(N)N}{f'(N)} - \varphi\right)\frac{f(N)}{f'(N)N}\frac{Y}{Y+\Theta} - \frac{1+g_y}{1+g_y-h}\right]$$
$$\Omega_3 \equiv \frac{h}{1+g_y-h}$$

Hence, potential output growth in this economy is defined as,

$$\Delta y_t^p = \theta_1 \Delta y_{t-1}^p + \theta_0 \varepsilon_{at} = \frac{\Omega_3}{\Omega_2}\Delta y_{t-1}^p + \frac{\Omega_1}{\Omega_2}\varepsilon_{at}$$

Given that the state variables of the model did not change with respect to the baseline version, the second equation of the SVAR does not change (see proposition 1).

$$\mu_t^w = \gamma_a \varepsilon_{at} + \gamma_\xi \xi_t + \gamma_\mu \mu_{t-1}^w + \gamma_y \Delta y_{t-1} \tag{B.2}$$

Now using (B.1) and (B.2) we can construct the following SVAR,

$$\begin{bmatrix} \Delta y_t \\ \mu_t^w \end{bmatrix} = \mathbf{B} \begin{bmatrix} \Delta y_{t-1} \\ \mu_{t-1}^w \end{bmatrix} + \mathbf{C} \begin{bmatrix} \varepsilon_{at} \\ \xi_t \end{bmatrix} \tag{B.3}$$

where,

$$\mathbf{B} \equiv \begin{bmatrix} \frac{\Omega_3-\gamma_y}{\Omega_2} & \frac{1-\gamma_\mu}{\Omega_2} \\ \gamma_y & \gamma_\mu \end{bmatrix} \qquad \mathbf{C} \equiv \begin{bmatrix} \frac{\Omega_1-\gamma_a}{\Omega_2} & \frac{-\gamma_\xi}{\Omega_2} \\ \gamma_a & \gamma_\xi \end{bmatrix}$$

Finally, using the matrices above it is easy to check that,

$$\theta_0 = c_{11} - \frac{c_{21}c_{12}}{c_{22}} \qquad \theta_1 = b_{11} - \frac{b_{21}c_{12}}{c_{22}}$$

2. *Using other preferences consistent with a BGP, with or without external habits*

Being more general we can assume that the instantaneous utility function is given by $U(C_t, \bar{C}_{t-1}, N_t)$, where U is consistent with a BGP. The relevant part of the model that changes is the marginal utility of consumption and the disutility of labor. Now we have,

$$mu_t = \frac{\bar{U}_{11}C}{\bar{U}_1}c_t + \frac{\bar{U}_{12}C}{\bar{U}_1}c_{t-1} + \frac{\bar{U}_{13}N}{\bar{U}_1}n_t$$

In this case we get the result,

$$\mu_t^w - \mu_{t-1}^w = \Omega_1\varepsilon_{at} - \Omega_2\Delta y_t + \Omega_3\Delta y_{t-1}$$

where

$$\begin{aligned}
\Omega_1 &\equiv \frac{1 + \frac{\bar{U}_{33}N}{\bar{U}_3} + \frac{\bar{U}_{13}N}{\bar{U}_1}}{1-\alpha}\sigma_a \\
\Omega_2 &\equiv \left[\frac{\frac{\bar{U}_{33}N}{\bar{U}_3} + \frac{\bar{U}_{13}N}{\bar{U}_1} - \alpha}{1-\alpha} + \frac{\bar{U}_{31}C}{\bar{U}_3} + \frac{\bar{U}_{11}C}{\bar{U}_1}\right] \\
\Omega_3 &\equiv -\left[\frac{\bar{U}_{32}C}{\bar{U}_3} + \frac{\bar{U}_{12}C}{\bar{U}_1}\right]
\end{aligned}$$

Hence, potential output growth in this economy is defined as,

$$\Delta y_t^p = \theta_1\Delta y_{t-1}^p + \theta_0\varepsilon_{at} = \frac{\Omega_3}{\Omega_2}\Delta y_{t-1}^p + \frac{\Omega_1}{\Omega_2}\varepsilon_{at}$$

The second equation in the SVAR does not change because there are not more state variables. Hence the SVAR is described by (B.3) and, as in the case where we changed the production function, the strategy to compute potential GDP does not change.

3. *Adding wage markup shocks.* Adding wage markup shocks incorporates an exogenous shock that directly affects the wage Phillips curve. Denoting with $\sigma_\mu\varepsilon_{\mu t}$ the new shock, Phillips curve is given by,

$$\pi_t^w = -\kappa^w\mu_t^w + \beta\mathbb{E}_t\pi_{t+1}^w + \sigma_\mu\varepsilon_{\mu t}$$

As a consequence, the only SVAR equation that is modified is the one that relates the wage markup with the state variables and shocks in the model. This equation is now,

$$\mu_t^w = \gamma_a \varepsilon_{at} + \gamma_z z_t + \gamma_\nu \nu_t + \gamma_\mu \varepsilon_{\mu t} + \gamma_\mu \mu_{t-1}^w + \gamma_y \Delta y_{t-1}$$

where the γ's are function of deep parameters of the model. However, this new shock only modifies the definition of the linear combination of shocks different from the TFP shock. We can now define $\xi_t = \frac{\gamma_z}{\sqrt{\gamma_z^2+\gamma_\nu^2+\gamma_\mu^2}} z_t + \frac{\gamma_\nu}{\sqrt{\gamma_z^2+\gamma_\nu^2+\gamma_\mu^2}} \nu_t + \frac{\gamma_\mu}{\sqrt{\gamma_z^2+\gamma_\nu^2+\gamma_\mu^2}} \varepsilon_{\mu t}$ and $\gamma_\xi = \sqrt{\gamma_z^2 + \gamma_\nu^2 + \gamma_\mu^2}$ then,

$$\mu_t^w = \gamma_a \varepsilon_{at} + \gamma_\xi \xi_t + \gamma_\mu \mu_{t-1}^w + \gamma_y \Delta y_{t-1} \tag{B.4}$$

The method to estimate potential GDP is unchanged after this modification in the definition of ξ.

4. *Other monetary policy rules.* As it is clear from the derivation of the baseline method, the monetary policy rule does not directly enter the SVAR. As a consequence, the effect of changing the monetary policy rule in the SVAR is indirect. In particular, changing the monetary policy rule affects the γ's in the second equation of the SVAR,

$$\mu_t^w = \gamma_a \varepsilon_{at} + \gamma_\xi \xi_t + \gamma_\mu \mu_{t-1}^w + \gamma_y \Delta y_{t-1}$$

However, this change does not affect the equations and, therefore, the method is not modified.

5. *Other expectation formation assumptions as long as they do not add more state variables.* Like the case of other monetary policy rules, expectations do not enter directly in the SVAR. As before, how agents forecast the future is embedded in the γ's in the equation

$$\mu_t^w = \gamma_a \varepsilon_{at} + \gamma_\xi \xi_t + \gamma_\mu \mu_{t-1}^w + \gamma_y \Delta y_{t-1}$$

Hence, as long as there is not a change in the set of state variables in the model, any change in expectation formation will only change the value of γ's without modifying the method to estimate potential GDP.

6. *Adding capital utilization as in Smets and Wouters (2007) and assuming capital in fixed*

supply (growing in a BGP)

This modification implies more substantial changes in the model than the previous cases. We start describing the key features that are changed in the model, and then show that the method to estimate potential output does not change. As before households maximize the utility,

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t Z_t \left[\log\left(C_t - h\bar{C}_{t-1}\right) - \chi \int_0^1 \frac{N_t(i)^{1+\varphi}}{1+\varphi} di \right]$$

However, they now decide how much capital services to lend to intermediate good producers. Capital services are defined by the multiplication between capital utilization U_t and physical capital $\bar{K}_t$ which is exogenously given (growing in a BGP). The budget constraint the household faces is the following. Now the household lends capital services $u_t\bar{K}_t$ to firms at a rate R_{kt}. $\gamma(U_t)$ represents capital utilization costs as in Smets and Wouters (2007).

$$P_t C_t + B_{t+1} = \int_0^1 W_t(i) N_t(i) di + R_{t-1} B_t + R_{kt} U_t \bar{K}_t + T_t - \gamma(U_t) P_t \bar{K}_t$$

After loglinearizing, detrending and some algebra we get,

$$\mu_t^w - \mu_{t-1}^w = -\Omega_2 \Delta y_t + \Omega_3 \Delta y_{t-1} + \Omega_1 \varepsilon_{at} \tag{B.5}$$

where,

$$\begin{aligned}
\Omega_1 &\equiv \frac{1+\varphi}{1-\alpha}\sigma_a \\
\Omega_2 &\equiv \left(\frac{1+g_y}{1+g-h}\Gamma + \frac{1+\varphi-(1-\alpha)\Theta}{(1-\alpha)\Theta} \right) \\
\Omega_3 &\equiv \frac{h}{1+g_y-h}\Gamma \\
\Gamma &\equiv \frac{Y}{C} + \frac{\gamma' \bar{K}/C}{(\gamma''+\gamma'(1-\alpha))\Theta} \\
\Theta &\equiv \frac{\gamma''+\gamma'(1-\alpha)+\alpha}{\gamma''+\gamma'(1-\alpha)}
\end{aligned}$$

and γ' and γ'' represent the first and second derivatives of the utilization cost function $\gamma(U_t)$ evaluated at steady state. Hence, potential output growth in this economy is defined as,

$$\Delta y_t^p = \theta_1 \Delta y_{t-1}^p + \theta_0 \varepsilon_{at} = \frac{\Omega_3}{\Omega_2} \Delta y_{t-1}^p + \frac{\Omega_1}{\Omega_2} \varepsilon_{at}$$

The second equation in the SVAR does not change because there are not more state variables. Hence the SVAR is described by (B.3) with new definitions for the Ω's and, as in the case where we changed the production function, the strategy to compute potential GDP remains the same.

C Proof of proposition 3

After detrending and loglinearizing, the two key equations in the model with price rigidities are,

$$\mu_t^w - \mu_{t-1}^w - \pi_t^w + \pi_t = -\left(\frac{1+g_y}{1+g_y-h} + \frac{\varphi}{1-\alpha}\right) \Delta y_t + \frac{h}{1+g_y-h} \Delta y_{t-1} + \frac{\varphi}{1-\alpha} \sigma_a \varepsilon_{at} \quad \text{(C.1)}$$

$$\frac{1}{1-\alpha} \sigma_a \varepsilon_{at} - \frac{\alpha}{1-\alpha} \Delta y_t = \pi_t^w - \pi_t + \mu_t^p - \mu_{t-1}^p \quad \text{(C.2)}$$

Combining (C.1) and (C.2) we can get the first equation in the system (remember $\tau_t = \mu_t^p + \mu_t^w$),

$$\tau_t = \mu_{t-1}^w + \mu_{t-1}^p - \left(\frac{1+g_y}{1+g_y-h} + \frac{\alpha+\varphi}{1-\alpha}\right) \Delta y_t + \frac{h}{1+g_y-h} \Delta y_{t-1} + \frac{1+\varphi}{1-\alpha} \sigma_a \varepsilon_{at} \quad \text{(C.3)}$$

The second equation comes from the fact that the labor wedge τ_t is a function of the state variables and shocks. Hence,

$$\tau_t = \eta_a \varepsilon_{at} + \eta_\xi \xi_t + \eta_w \mu_{t-1}^w + \eta_p \mu_{t-1}^p + \eta_y \Delta y_{t-1} \quad \text{(C.4)}$$

where ξ_t is just a linear combination of preference shocks z_t and monetary policy shocks ν_t. Equations (C.3) and (C.4) form the following system,

$$\begin{bmatrix} \Delta y_t \\ \tau_t \end{bmatrix} = \mathbf{B} \begin{bmatrix} \Delta y_{t-1} \\ \mu^w_{t-1} \\ \mu^p_{t-1} \end{bmatrix} + \mathbf{C} \begin{bmatrix} \varepsilon_{at} \\ \xi_t \end{bmatrix}$$

where

$$\mathbf{B} = \begin{bmatrix} \frac{\frac{h}{1+g_y-h}-\eta_y}{\frac{1+g_y}{1+g_y-h}+\frac{\alpha+\varphi}{1-\alpha}} & \frac{1-\eta_w}{\frac{1+g_y}{1+g_y-h}+\frac{\alpha+\varphi}{1-\alpha}} & \frac{1-\eta_p}{\frac{1+g_y}{1+g_y-h}+\frac{\alpha+\varphi}{1-\alpha}} \\ \eta_y & \eta_w & \eta_p \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} \frac{\frac{1+\varphi}{1-\alpha}\sigma_a-\eta_a}{\frac{1+g_y}{1+g_y-h}+\frac{\alpha+\varphi}{1-\alpha}} & \frac{-\eta_\xi}{\frac{1+g_y}{1+g_y-h}+\frac{\alpha+\varphi}{1-\alpha}} \\ \eta_a & \eta_\xi \end{bmatrix}$$

Finally, using the matrices above it is easy to check that,

$$\theta_0 = c_{11} - \frac{c_{21}c_{12}}{c_{22}} \qquad \theta_1 = b_{11} - \frac{b_{21}c_{12}}{c_{22}}$$

Also, having the matrices **B** and **C** it is easy to get the structural shocks ξ_t and ε_{at},

$$\mathbf{C}^{-1} \left\{ \begin{bmatrix} \Delta y_t \\ \tau_t \end{bmatrix} - \mathbf{B} \begin{bmatrix} \Delta y_{t-1} \\ \mu^w_{t-1} \\ \mu^p_{t-1} \end{bmatrix} \right\} = \begin{bmatrix} \varepsilon_{at} \\ \xi_t \end{bmatrix} \tag{C.5}$$

for all t.

D Data Sources

The data for GDP growth (Δy_t^{data}), wage markups ($\mu_t^{w,data}$), price markups ($\mu_t^{p,data}$) and government spending growth (Δg_t^{data}) were constructed using the following equations,

$$
\begin{aligned}
\Delta y_t^{data} &= \log\left(Y_t^{data}\right) - \log\left(Y_{t-1}^{data}\right) \\
\mu_t^{w,data} &= \log\left(\frac{W_t^{data}}{P_t^{data}}\right) - \log\left(N_t^{data}\right) - \log\left(C_t^{data}\right) \\
\mu_t^{p,data} &= \log\left(\frac{Y_t^{data}}{POP_t}\right) - \log\left(N_t^{data}\right) - \log\left(\frac{W_t^{data}}{P_t^{data}}\right) \\
\Delta g_t^{data} &= \log\left(G_t^{data}\right) - \log\left(G_{t-1}^{data}\right)
\end{aligned}
$$

where Y_t^{data} is real GDP (GDPC1), W_t^{data} is the nonfarm business sector compensation per hour (COMPNFB), P_t^{data} is the GDP deflator (GDPDEF). Moreover, N_t^{data} = HOANBS/CNP16OV where HOANBS is BLS total hours in the nonfarm business sector and CNP16OV is total population over 16 years old. Finally, C_t^{data} denotes Real personal consumption expenditures per capita (A794RX0Q048SBEA), POP_t represents total population over 16 years old (CNP16OV), and G_t^{data} is real government consumption expenditures and gross investment (GCEC1). All these series were downloaded from FRED [Link here].

We use other series either as instruments in the SVAR estimation or for comparison with our results. CBO's potential output estimates were obtained from CBO's website: Potential GDP and Underlying Inputs (August 2019 version) [Link here] updated with 10-Year Economic Projections (February 2021 version) [Link here] for the recent estimates. Fernald's TFP can be downloaded from the San Francisco Fed website (May 2020 version) [Link here]. For Romer and Romer monetary policy shock series, we use the one updated by Coibion et al. (2017a). Lastly, unemployment series (UNRATE) were downloaded from FRED.

References

Andrés, J., J. D. López-Salido, and E. Nelson (2005): "Sticky-price models and the natural rate hypothesis," *Journal of Monetary Economics*, 52, 1025 – 1053.

Anzoategui, D., D. Comin, M. Gertler, and J. Martinez (2019): "Endogenous Technology Adoption and R&D as Sources of Business Cycle Persistence," *American Economic Journal: Macroeconomics*, 11, 67–110.

Auerbach, A. J. and Y. Gorodnichenko (2012): "Measuring the Output Responses to Fiscal Policy," *American Economic Journal: Economic Policy*, 4, 1–27.

Basu, S. and J. G. Fernald (2009): "What do we know and not know about potential output?" Working Paper Series 2009-05, Federal Reserve Bank of San Francisco.

Basu, S., J. G. Fernald, and M. S. Kimball (2006): "Are Technology Improvements Contractionary?" *American Economic Review*, 96, 1418–1448.

Beraja, M. (2019): "Counterfactual Equivalence in Macroeconomics," Working paper, MIT.

Beveridge, S. and C. R. Nelson (1981): "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'," *Journal of Monetary Economics*, 7, 151 – 174.

Bils, M., P. J. Klenow, and B. A. Malin (2018): "Resurrecting the Role of the Product Market Wedge in Recessions," *American Economic Review*, 108, 1118–46.

Blanchard, O., E. Cerutti, and L. Summers (2015): "Inflation and Activity – Two Explorations and their Monetary Policy Implications," Working Paper 21726, National Bureau of Economic Research.

Blanchard, O. and R. Perotti (2002): "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *The Quarterly Journal of Economics*, 117, 1329–1368.

Blanchard, O. J. and D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *The American Economic Review*, 79, 655–673.

Brinca, P., J. D. Duarte, and M. Faria-e Castro (2020): "Measuring Labor Supply and Demand Shocks during COVID-19," Working paper.

Calvo, G. A. (1983): "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics*, 12, 383 – 398.

Canova, F. and L. Sala (2009): "Back to square one: Identification issues in DSGE models," *Journal of Monetary Economics*, 56, 431 – 449.

Cerra, V., A. Fatas, and S. C. Saxena (2020): "Hysteresis and Business Cycles," Working Paper 20/73, International Monetary Fund.

Cerra, V. and S. C. Saxena (2008): "Growth Dynamics: The Myth of Economic Recovery," *American Economic Review*, 98, 439–57.

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

——— (2009): "New Keynesian Models: Not Yet Useful for Policy Analysis," *American Economic Journal: Macroeconomics*, 1, 242–66.

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

Cochrane, D. and G. H. Orcutt (1949): "Application of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms," *Journal of the American Statistical Association*, 44, 32–61.

Coibion, O., Y. Gorodnichenko, L. Kueng, and J. Silvia (2017a): "Innocent Bystanders? Monetary policy and inequality," *Journal of Monetary Economics*, 88, 70–89.

Coibion, O., Y. Gorodnichenko, and M. Ulate (2017b): "The Cyclical Sensitivity in Estimates of Potential Output," Working Paper 23580, National Bureau of Economic Research.

Comin, D. and M. Gertler (2006): "Medium-Term Business Cycles," *American Economic Review*, 96, 523–551.

Comin, D. A., J. Q. Gonzalez, T. G. Schmitz, and A. Trigari (2020): "Measuring TFP: The Role of Profits, Adjustment Costs, and Capacity Utilization," Working Paper 28008, National Bureau of Economic Research.

Edge, R. M., M. T. Kiley, and J.-P. Laforte (2008): "Natural rate measures in an estimated DSGE model of the U.S. economy," *Journal of Economic Dynamics and Control*, 32, 2512 – 2535.

Erceg, C. J., D. W. Henderson, and A. T. Levin (2000): "Optimal monetary policy with staggered wage and price contracts," *Journal of Monetary Economics*, 46, 281–313.

Fatás, A. (2000): "Endogenous growth and stochastic trends," *Journal of Monetary Economics*, 45, 107 – 128.

Fatás, A. and L. H. Summers (2018): "The permanent effects of fiscal consolidations," *Journal of International Economics*, 112, 238 – 250.

Fernald, J. G. (2007): "Trend breaks, long-run restrictions, and contractionary technology improvements," *Journal of Monetary Economics*, 54, 2467 – 2485.

——— (2012): "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," FRBSF Working Paper 2012-19, National Bureau of Economic Research.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, T. J. Sargent, and M. W. Watson (2007): "ABCs (and Ds) of Understanding VARs," *American Economic Review*, 97, 1021–1026.

Francis, N. and V. A. Ramey (2005): "Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited," *Journal of Monetary Economics*, 52, 1379 – 1399.

Furlanetto, F., Ø. Robstad, P. Ulvedal, and A. Lepetit (2020): "Estimating hysteresis effects," Working paper 13/20, Norges Bank.

Galí, J. (1999): "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review*, 89, 249–271.

——— (2015): *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications Second edition*, no. 10495 in Economics Books, Princeton University Press.

Galí, J., M. Gertler, and J. D. López-Salido (2007): "Markups, Gaps, and the Welfare Costs of Business Fluctuations," *The Review of Economics and Statistics*, 89, 44–59.

Galí, J., F. Smets, and R. Wouters (2012): "Unemployment in an Estimated New Keynesian Model," *NBER Macroeconomics Annual*, 26, 329 – 360.

Gordon, R. J. (2014): "A New Method of Estimating Potential Real GDP Growth: Implications for the Labor Market and the Debt/GDP Ratio," NBER Working Papers 20423, National Bureau of Economic Research, Inc.

Hall, R. E. (2009): "Reconciling Cyclical Movements in the Marginal Value of Time and the Marginal Product of Labor," *Journal of Political Economy*, 117, 281–323.

Hodrick, R. J. and E. Prescott (1981): "Post-War U.S. Business Cycles: An Empirical Investigation," Discussion Papers 451, Northwestern University, Center for Mathematical Studies in Economics and Management Science.

Jordà, O. (2005): "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, 95, 161–182.

Jordà, O., M. Schularick, and A. M. Taylor (2013): "When Credit Bites Back," *Journal of Money, Credit and Banking*, 45, 3–28.

Jordà, O., S. R. Singh, and A. M. Taylor (2020): "The Long-Run Effects of Monetary Policy," Working Paper 26666, National Bureau of Economic Research.

Justiniano, A., G. E. Primiceri, and A. Tambalotti (2013): "Is There a Trade-Off between Inflation and Output Stabilization?" *American Economic Journal: Macroeconomics*, 5, 1–31.

Karabarbounis, L. (2014): "The labor wedge: MRS vs. MPN," *Review of Economic Dynamics*, 17, 206 – 223.

Lucas, R. J. (1976): "Econometric policy evaluation: A critique," *Carnegie-Rochester Conference Series on Public Policy*, 1, 19–46.

Mishkin, F. S. (2007): "Estimating Potential Output," Speech at the Conference on Price Measurement for Monetary Policy, Federal Reserve Bank of Dallas, Dallas, Texas.

Neiss, K. S. and E. Nelson (2005): "Inflation Dynamics, Marginal Cost, and the Output Gap: Evidence from Three Countries," *Journal of Money, Credit and Banking*, 37, 1019–1045.

Nekarda, C. J. and V. A. Ramey (2020): "The Cyclical Behavior of the Price-Cost Markup," mimeo, UCSD.

Okun, M. A. (1962): "Potential GDP:Its Measurement and Significance," Working Paper 190, Cowles Foundation, Yale.

Ramey, V. (2016): "Chapter 2 - Macroeconomic Shocks and Their Propagation," Elsevier, vol. 2 of *Handbook of Macroeconomics*, 71 – 162.

Ramey, V. A. (2011): "Identifying Government Spending Shocks: It's all in the Timing*," *The Quarterly Journal of Economics*, 126, 1–50.

Ravenna, F. (2007): "Vector autoregressions and reduced form representations of DSGE models," *Journal of Monetary Economics*, 54, 2048 – 2064.

Reinhart, C. M. and K. S. Rogoff (2014): "Recovery from Financial Crises: Evidence from 100 Episodes," *American Economic Review*, 104, 50–55.

Romer, C. D. and D. H. Romer (2004): "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, 94, 1055–1084.

Sala, L., U. Söderström, and A. Trigari (2008): "Monetary policy under uncertainty in an estimated model with labor market frictions," *Journal of Monetary Economics*, 55, 983 – 1006.

SHACKLETON, R. (2018): "Estimating and Projecting Potential Output Using CBO's Forecasting Growth Model," Working Paper 2018-03, Congressional Budget Office.

SHIMER, R. (2009): "Convergence in Macroeconomics: The Labor Wedge," *American Economic Journal: Macroeconomics*, 1, 280–97.

SMETS, F. AND R. WOUTERS (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, 586–606.

STADLER, G. W. (1990): "Business Cycle Models with Endogenous Technology," *The American Economic Review*, 80, 763–778.

STIGLITZ, J. E. (1993): "Endogenous Growth and Cycles," Working Paper 4286, National Bureau of Economic Research.

STOCK, J. H. (2008): "Lecture 7 - Recent Developments in Structural VAR Modeling," Lecture, NBER Summer Institute Econometrics Series.

WOODFORD, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.