Re-examining the Output Gap Restriction and Role of Aggregate Supply in Macroeconomic Fluctuations

James Dean*
Western Carolina University

Scott Schuh†
West Virginia University

March 20, 2025

Preliminary and Incomplete.

Abstract

This paper re-examines the output gap restriction in small macro models and finds an overlooked independent role for high-frequency fluctuations in potential output, especially in the Phillips curve and Taylor Rule. Surprisingly, the data reject the conventional (but untested) symmetric restriction on actual and potential output prior to unconventional monetary policy (1960-2007). Historically, shocks to long-run aggregate supply had low variance and modest economic impact on model dynamics. Recent literature questions whether the COVID-19 Pandemic recession was an unusual shock to long-run aggregate supply. Out-of-sample simulations of the unrestricted model suggest the COVID recession may be better explained as a large, highly transitory shock to potential output than a traditional aggregate demand shock.

Keywords: Output gap, aggregate supply, COVID recession

JEL Classifications: E32, E52, C50, C30

Email: jdean@wcu.edu
†*Email*: scott.schuh@mail.wvu.edu

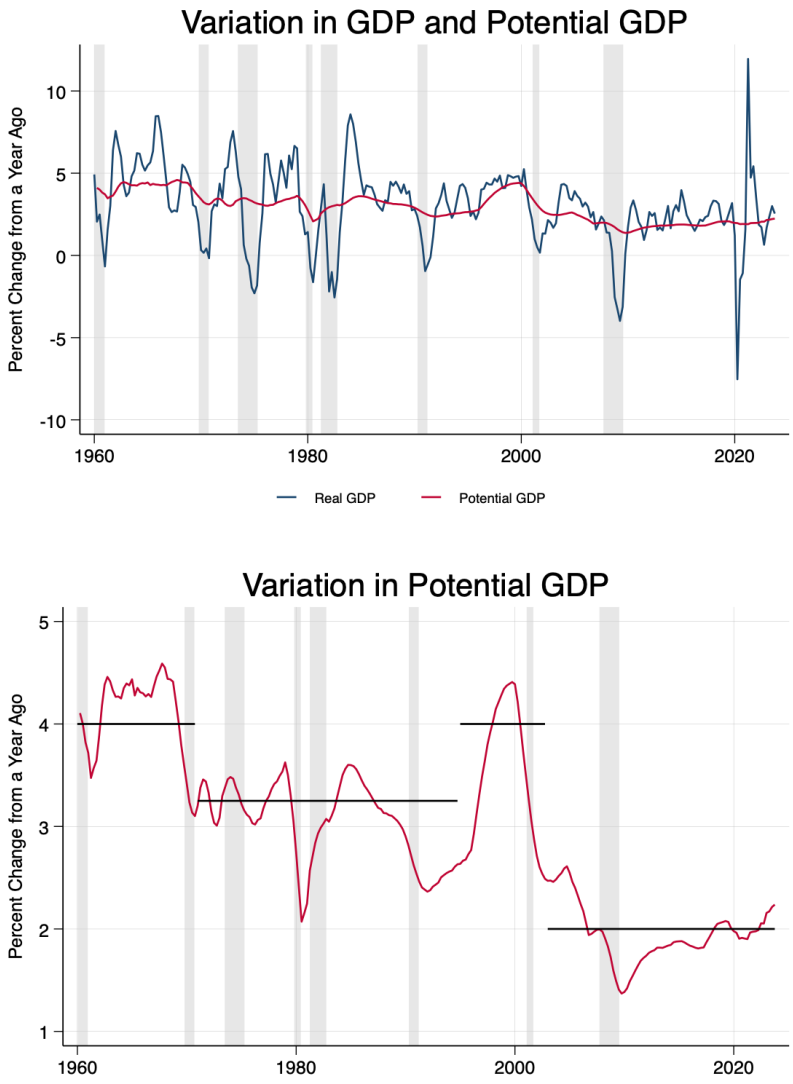
1 Introduction

The output gap concept is central to macroeconomic analysis and embraced in some form by most macroeconomists. The gap between actual and potential real GDP (output) summarizes aggregate utilization in output markets and is closely related to the unemployment rate gap via Okun's Law. Utilization gaps are viewed as the driving force in Phillips curve models of inflation dynamics and a useful measure of the state of dynamic adjustment to long-run steady state during business cycles. The output gap specification implies that actual and potential output have symmetric effects on other variables in the model.

Practical measurement of the output gap, on the other hand, is widely recognized to be challenging and imperfect. Measurement of actual GDP is relatively successful, albeit subject to non-trivial measurement error [CITE] and potentially large revisions in real-time [CITE]. Measurement of potential GDP adds an additional layer of complication and uncertainty related to estimation of its trend growth (Arnold, 2009). The conventional view is that potential output growth is constant, or approximately so, thus leaving actual output as the main driver of business cycle fluctuations in the gap.

Figure 1 illustrates the conventional view and contrasting reality. The top panel, which plots the growth rates (4-quarter percent changes) of actual and potential real GDP, illustrates that potential GDP growth is *relatively* constant compared with actual GDP. Careful inspection reveals that potential GDP growth is not constant but fluctuates at longer-run (decadal) frequencies, as previously documented (Benati, 2007; Fernald, 2015; Kahn & Rich, 2007). The bottom panel zooms in on potential GDP growth and adds black lines for the conventional trend breaks (1973, 1995, and 2005) to highlight a largely overlooked fact: *potential output growth fluctuates far more in the short run than previously recognized.* Potential output growth often fluctuates 0.5 percentage point, and sometime more than 1 percentage point, *within* a business cycle.

Figure 1: Actual versus Potential Real GDP Growth



Perhaps for this reason, the output gap restriction of equal and opposite roles for actual and potential output in business cycle fluctuations has not been tested. Traditionally, the output gap appears in macro models with a single coefficient: $\alpha(y_t – y_t^*)$. This specification implicitly assumes that actual (y_t) and potential (y_t^*) output have equal and opposite-signed coefficients: $\alpha_y = -\alpha_{y*}$. Trend growth in potential GDP may fluctuate in the long or medium run *across* business cycles, but the output gap restriction rules out an independent role for quarter-to-quarter fluctuations in potential GDP *within* business cycles. Thus, the output

gap restriction imposes the traditional macro view of orthogonal business cycles (output gap) and growth (trend potential output).

Some contemporary macro modeling approaches relax the strict orthogonal growth-cycle decomposition and admit covariance between the cycle and growth.[1] One simple way to introduce this covariance is by relaxing the output gap restriction, $\alpha_y \neq \alpha_{y*}$, which induces three important changes the the model. First, y_t and y_t^* have independent roles in the business cycle and effects on other endogenous variables.[2] Second, y_t^* becomes an endogenous variable that requires structure to govern its relationships with y_t in the long and short run (e.g., co-integration and error correction). Third, properly specified general equilibrium frameworks require a structural error in potential output, ε_t^{y*}. Together, these changes enable macro models to entertain high-frequency (quarterly) shifts in y_t^*, or long-run aggregate supply (LRAS), attributable to endogenous dynamics or structural shocks. Real business cycle models also have this property but focus on fluctuations in actual TFP, A_t, rather than potential output, $y_t^* = A_t^* f(L_t^*, K_t^*)$.

The unprecedented COVID-19 Pandemic raises intriguing questions about identification of the shock(s), propagation, and amplification of the shortest (Feb to April, 2020) and perhaps most unusual U.S. recession. In 2020:H1, real GDP (the real output gap) declined about 8 percent (percentage points) followed by a nearly equal rebound shortly afterward. Most observers interpreted the COVID pandemic as a traditional shock to aggregate demand (AD), or y_t, albeit for different reasons. Rather than reducing consumer expenditures due to layoffs, unemployment, and lower income, consumers were unable (or unwilling) to buy goods and services in public. Highly stimulative monetary and fiscal policies were implemented as the traditional prescription for the COVID recession.

However, recent literature questions whether the unexpected COVID-19 pandemic and recession might be explained as a change in LRAS rather than aggregate demand. Campbell

[1] For examples, RBC and VEC models...(brief lit review forthcoming here)

[2] In theory, y_t and y_t^* could have independent roles growth as well, but that feature would require major deviations from contemporary equilibrium-based approaches to macro.

and Mankiw (2024) reminder, "Disentangling shocks to supply and demand is nearly impossible in real time," is not debatable. But Reis (2022) argues, "Many of these [macro COVID] shocks can just as well be interpreted as shocks to the potential level of output." Guerrieri et al. (2022) use a two-sector model to interpret COVID as a "Keynesian supply shock" that was "large, asymmetric, and transitory," which reduced income and ultimately worked as a decline in aggregate demand that was ultimately overcome by fiscal stimulus. Similarly, Baqaee and Farhi (2022) build a multi-sector model with inflationary supply shocks that become deflationary demand shocks that also inhibit the effectiveness of fiscal policy.[3] Small-scale macro models with an unrestricted output gap framework offer a cyclical role for potential output that identifies AD versus LRAS in a parsimonious way that can be explained in Principles of Macroeconomics.

This paper re-examines the traditional role of the output gap restriction in business cycles using small-scale VAR models with $z_t = [\pi_t \; y_t \; y_t^* \; i_t]'$ where π_t is inflation and i_t is the short-term interest rate. It shows that the output gap specification produces a restricted 3-variable VAR model (henceforth, VAR3=VAR4r) from an unrestricted 4-variable VAR model (VAR4u) conducts formal econometric tests of the VAR3 output gap restrictions. This analysis is extended to a 4-variable *structural* VAR model that imposes the output gap restriction (SVAR4r) or not (SVAR4u) to corroborate the reduced-form VAR analysis. The econometric models impose long-run co-integrating vectors: 1) between y_t and y_t^*; and 2) between $\imath_t$ and π_t (Fisher equation). Estimation is 1960:Q1 to 2007:Q4 to avoid complications of unconventional monetary policy (e.g., Mavroeidis, 2021) and demonstrate a cyclical role for LRAS before the Global Financial Crisis. Model dynamics (impulse responses) are compared and the estimated SVAR4u model is simulated to evaluate the COVID recession.

Perhaps surprisingly, the data prior to 2008 reject the output gap restriction statistically significantly. In VAR4u model, the rejection primarily occurs among the block of coefficients on y_t and y_t^* in the first lagged coefficient matrix ($B_1 z_{t-1}$). In SVAR4 models, the rejection

[3]Several other papers offer related models... TO BE COMPLETED.

primary occurs in the output gap restrictions of the impact matrix coefficients ($B_0 z_t$) for the Phillips curve and Taylor Rule (inflation and interest rate equations, respectively). Notably, inflation and interest rates respond more strongly to movements in potential output than actual output. While these differences are theoretically plausible, they represent an unfamiliar perspective when applied to business cycle analysis.

The impact on model dynamics (impulse responses) of relaxing the output gap restriction is economically modest, however. Reduced-form VAR impulse responses (VAR4r=VAR3 and VAR4u) are not economically different except for the response of inflation to a shock in potential output. The restricted model (VAR4r=VAR3) predicts inflation returns to its steady state after a shock to potential output, but the unrestricted model (VAR4u) predicts a positive response of about one-half percent percentage point for 20 quarters. In SVAR models, dynamic differences between SVAR4r and SVAR4u are notably larger, mainly in the responses to shocks in potential output. The unrestricted model (SVAR4u) also has larger and more persistent effects of output shocks on interest rates and especially inflation.

Out-of-sample (2020-2024) simulations of the pre-2008 estimated SVARu model provide thought-providing evidence that the COVID recession may be better characterized as a very large but transitory shock to LRAS (ε_{y^*}) rather than AD (ε_y). The model simulation applies structural shocks that generate data-consistent paths for y_t and y_t^* data during the early stages of the COVID recession (2020:Q1 through 2020:Q3) then allow the dynamics to be governed by the estimated model. These simulations show that the LRAS shocks fit the actual observed data during 2020:Q4 through 2023:Q4 more closely than the data-consistent shocks to AD. However, neither shock adequately explains the actual macroeconomic responses to the COVID pandemic, most likely because the model omits the massive fiscal stimulus (including money growth) needed to explain the U.S. inflation surge (Campbell & Mankiw, 2025).

The results in this paper have potentially important implications for monetary and fiscal policy. Estimates and dynamics the unrestricted model suggest that it may be important for

policy makers to distinguish more carefully between the structure and shocks in actual and potential output at high frequency. Not only does potential output have distinctly different effects on the endogenous variables in the macroeconomy in the long and short run, but also distinctly different effects on policy instruments via policy rules. This finding is not new to the COVID pandemic. Historical episodes also support this conclusion, from the productivity surge during the dot-com era to the subsequent productivity decline following the Global Financial Crisis. Policy rules that respond only to the output gap without differentiating the individual behavior of its components in the rule and real economy may be suboptimal.

2 Previous Literature

The modern concept of potential output was first formally introduced by Okun (1970), who linked it to the relationship between output and unemployment, now widely known as Okun's Law. Okun's early work laid the foundation for measuring the output gap, emphasizing the role of labor market conditions in defining economic capacity. Building on this, Perry (1978) formally integrated labor utilization into estimates of potential GDP, emphasizing the role of unemployment benchmarks in determining the output gap.

Clark (1983) and Clark (1987) decomposed U.S. economic activity into cyclical and trend components, suggesting that at least half of the quarterly variations in output can be attributed to persistent cyclical effects rather than smooth, deterministic trends. Kuttner (1994) estimates potential GDP by treating it as a latent variable and using the Kalman filter in conjunction with an aggregate supply relationship. Meanwhile, Campbell and Mankiw (1987) showed that economic shocks can have persistent effects on output, thus questioning the conventional understanding of business cycles. Finally, Boschen and Mills (1990) examined the role that higher-frequency variations in potential GDP have on policy setting, emphasizing the impact the measurement of potential GDP has on policy decisions.

It is widely recognized that potential output growth rates have shifted *across* business

cycles, mainly due to breaks in trend productivity. Benati (2007) examined the role of structural breaks in labor productivity, showing that long-term changes in productivity trends can significantly alter potential output estimates. Similarly, Kahn and Rich (2007) employed regime-switching models to identify shifts in trend productivity, reinforcing the idea that potential output evolves over time rather than following a fixed trajectory.

While understanding variation in potential output is essential, estimating it accurately remains a persistent challenge. Traditional methods often conflate cyclical fluctuations with structural changes, leading to biased estimates. Coibion et al. (2017) examine common filtering techniques, showing how they frequently misclassify transitory downturns as permanent declines in potential GDP, particularly following recessions. This misclassification can result in overly pessimistic assessments of economic capacity and misguided policy responses. On the other hand, Martin et al. (2015) analyzed data across multiple advanced economies and found that deep recessions tend to result in downward revisions of potential output rather than rapid post-recession recoveries.

Real-time estimation of potential output presents a unique set of challenges due to the inherent uncertainty in preliminary data and the need for frequent revisions. Arnold (2009) focuses on the instability of real-time estimates, demonstrating how early readings of potential output are often significantly revised as more data becomes available. These revisions can lead to misinterpretations of the economy's underlying strength or weakness, complicating efforts to discern true structural changes from cyclical noise. Arnold emphasizes that the difficulty lies not only in data quality but also in methodological choices, as models often struggle to filter out short-term fluctuations without distorting longer-term trends. This highlights the precarious balance between responsiveness and stability in potential output estimation.

These challenges are further illustrated in Shackleton (2018), who provides an in-depth look at the Congressional Budget Office's (CBO) approach to estimating potential output. The CBO's methodology, which integrates labor force participation, capital accumulation,

and total factor productivity, aims to create a comprehensive measure of economic capacity. However, Shackleton points out that the CBO's reliance on long-run projections makes it less sensitive to short-term disruptions, leading to potential lags in recognizing real-time shifts. Despite ongoing refinements, significant gaps remain, particularly in accounting for unexpected economic shocks and rapid structural changes.

Taken as a whole, potential output simultaneously central to most macroeconomic research, yet fraught with both real-time and ex-post estimation challenges. Far from being a fixed anchor, potential output is dynamic, influenced by both long-run structural trends and short-term cyclical forces.

Impact of Covid- For later? (Ferroni et al., 2024), (Guerrieri et al., 2022), (Hall & Kudlyak, 2023), (Crump et al., 2024), (Fernald & Li, 2022) Ricardo Reis, Greg Mankiw, others have made short comments about AS in papers without specific models of COVID recession.

3 Benchmark Macro Models

This section describes several small-scale benchmark macroeconomic models used in the analysis. A reduced-form VAR allows straightforward econometric testing of the output gap restriction and initial assessment of the role of potential output as an independent variable. A structural VAR and New Keynesian model are used to evaluate robustness of the results from the reduced-form VAR to the introduction of more structural identification and rational expectations.

3.1 Reduced-form VAR

The simplest reduced-form macro VAR model includes three variables (VAR3): y_t denotes actual real output (GDP), π_t inflation, and i_t the nominal short-term interest rate. Let

$\widetilde{y}_t = y_t - y^*$ denote the output gap, where y^* is real potential output (GDP*) or long-run aggregate supply (LRAS). The VAR3(1) model ($L = 1$ lags) is:

$$z_{3,t} = B_{3,1} z_{3,t-1} + \varepsilon_{3,t} \tag{1}$$

where $z_{3,t} = [\pi_t \ \widetilde{y}_t \ i_t]'$ and $\varepsilon_{3,t} = [\varepsilon_t^{\pi} \ \varepsilon_t^{\widetilde{y}} \ \varepsilon_t^{i}]'$ and

$$B_{3,1} = \begin{bmatrix} b_{\pi\pi} & b_{\pi\widetilde{y}_t} & b_{\pi i} \\ b_{y\pi} & b_{y\widetilde{y}} & b_{yi} \\ b_{i\pi} & b_{i\widetilde{y}} & b_{ii} \end{bmatrix}.$$

As ordered, the model can be identified structurally with the usual Cholesky decomposition to derive the impact matrix B_0^{-1} from $\varepsilon_{3,t}$. The ordering rationale is that inflation (the aggregate price) is stickier than output and monetary policy can respond more quickly to shocks than output or inflation.

The benchmark VAR3 model implicitly imposes two key assumptions, which are readily accepted but rarely tested.

Assumption 1 *Potential output is approximately constant during a business cycle: $y_t^* \approx y^*$.*

Although it is widely recognized that y^* varies at the medium- and long-run frequencies, Assumption 1 tends to hold at least implicitly *within* each cycle given the long-run nature of y^*. Put differently, macroeconomics typically do not view y^* as an important contributor to fluctuations within a business cycle. Thus, Assumption 1 is relaxed by allowing potential output to vary quarter-to-quarter as in the data: $y^* = y_t^*$.[4]

A second assumption is specifying y_t and y_t^* to appear as the output gap, $\widetilde{y}_t$, in the model. This assumption restricts components of the gap to have equal and opposite signs, which eliminates any role for independent fluctuations in y_t^* to contribute to the business cycle. To

[4]In practice, macro models do not explicitly impose Assumption 1, especially if they use pre-constructed data like that from the Congressional Budget Office in FRED (mnemonic GDPPOT). Most macroeconomic theory, however, omits a cyclical role for quarterly variation in y^*.

see this, consider the unrestricted reduced-form VAR model with endogenous y_t^* (VAR4u):

$$z_{4,t} = B_{4,1} z_{4,t-1} + \varepsilon_{4,t} \tag{2}$$

where $z_{4,t} = [\pi_t \; y_t \; y_t^* \; i_t]'$, $\varepsilon_{4,t} = [\varepsilon_t^{\pi} \; \varepsilon_t^{y} \; \varepsilon_t^{y^*} \; \varepsilon_t^{i}]'$ and

$$B_{4,1} = \begin{bmatrix} b_{\pi\pi} & b_{\pi y} & b_{\pi y^*} & b_{\pi i} \\ b_{y\pi} & b_{yy} & b_{yy^*} & b_{yi} \\ b_{y^*\pi} & b_{y^*y} & b_{y^*y^*} & b_{y^*i} \\ b_{i\pi} & b_{iy} & b_{iy^*} & b_{ii} \end{bmatrix}.$$

The VAR3 model with $(\widetilde{y}_t)$ imposes the following restrictions on VAR4u:

Assumption 2 *The output gap specification restricts coefficients in $B_{4,1}$ on y_t and y_t^* to have equal and opposite signs: $b_{\pi y} = -b_{y^*\pi}$; $b_{iy} = -b_{iy^*}$; and $(b_{yy} - b_{y^*y}) = -(b_{yy^*} - b_{y^*y^*})$.*

Assumption 2 has testable implications for analogous parameters of $B_{3,1}$ and $B_{4,1}$:

Corollary 1 *Imposing the output gap restriction implies equality of coefficients on π_t and i_t in VAR3 and VAR4u output equations: $b_{3,\widetilde{y}_t\pi} = (b_{4,y\pi} - b_{4,y^*\pi})$ and $b_{3,\widetilde{y}_t i} = (b_{4,yi} - b_{4,y^*i})$.*

Although the restrictions in Corollary 1 are not formally imposed in VAR3 or VAR4u, they should hold when restrictions in Assumption 2 are imposed. Assumption 2 and Corollary 1 show that VAR4u fully nests VAR3, but to our knowledge these restrictions have not been tested econometrically. Thus, we relax Assumptions 1 and 2 and test formally whether they are consistent with the data.

Two observations are warranted. First, VAR4u introduces a role for potential output (LRAS) to influence the macroeconomy over the business cycle *independently* from actual output (i.e., aggregate demand). Conventional macro theory rules out this effect by assumption; we ask whether the assumptions are consistent with the data. These conventional models assume that shocks to the output gap are due solely to shocks to actual output,

so shocks to potential output would be unimportant. On the contrary, our specification allows shocks to potential output to have heterogeneous effects compared to actual output shocks. Second, identification of VAR4u is challenging for many reasons. For examples, the model does not include standard determinants of long-run aggregate (technology and factors of production) or a theory of trends (deterministic or stochastic). One solution is a different ordering with $z_{4,t} = [y_t^* \pi_t \ y_t \ i_t]'$ to account for relatively slower adjustment of long-run variables like y_t^*.

3.2 Structural VAR

An alternative strategy for identification is to impose simultaneous equation restrictions in a structural VAR (SVAR) model without forward-looking expectations. Let $r_t = i_t - \pi_t$ denote the *ex post* real interest rate. Abstracting from constants, the simultaneous equations of a 3-variable SVAR that is consistent with a basic New Keynesian (NK) model without expectations can be written as:[5]

$$\pi_t = \alpha \widetilde{y}_t + \varepsilon_t^{\pi} \tag{3}$$

$$\widetilde{y}_t = -\gamma r_{t-1} + \varepsilon_t^{\widetilde{y}} \tag{4}$$

$$i_t = \phi_\pi i_t + \phi_{\widetilde{y}} \widetilde{y}_t + \varepsilon_t^{i} \tag{5}$$

where all parameters are assumed to be positive ($\gamma, \alpha, \phi_\pi, \phi_{\widetilde{y}} > 0$). Equation (3) represents the Phillips curve, (4) the IS curve, and (5) the Taylor Rule. In matrix form, the 3-variable SVAR model (SVAR3) is:

$$\begin{bmatrix} 1 & -\alpha & 0 \\ -\gamma & 1 & \gamma \\ -\phi_\pi & -\phi_{\widetilde{y}} & 1 \end{bmatrix} \begin{bmatrix} \pi_t \\ \widetilde{y}_t \\ i_t \end{bmatrix} = \begin{bmatrix} b_{\pi\pi} & b_{\pi\widetilde{y}} & b_{\pi i} \\ b_{y\pi} & b_{\widetilde{y}\widetilde{y}} & b_{yi} \\ b_{i\pi} & b_{i\widetilde{y}} & b_{ii} \end{bmatrix} \begin{bmatrix} \pi_{t-1} \\ \widetilde{y}_{t-1} \\ i_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^{\pi} \\ \varepsilon_t^{\widetilde{y}} \\ \varepsilon_t^{i} \end{bmatrix} . \tag{6}$$

[5] See **TBD** and **TBD** for expositions of the original 3-variable NK model that includes forward-looking (rational) expectations.

In addition to Assumptions 1 and 2, SVAR3 assumes additional restrictions on B_0:

Assumption 3 *The output gap:* $\alpha_y = -\alpha_{y^*}$, $\phi_y = -\phi_{y^*}$, *and* $\delta_{yy^*} = 1$.

Assumption 4 *The real interest rate:* $\gamma_{yi} = -\gamma_{y\pi}$.

Most macroeconomists readily accept, but rarely test, Assumptions 3 and 4 as well.

Relaxing Assumptions 1 -4 and Corollary 1 yields an unrestricted 4-variable simultaneous equations model with endogenous y_t^*:

$$\pi_t = \alpha_y y_t + \alpha_{y^*} y_t^* + \varepsilon_t^{\pi} \tag{7}$$

$$y_t = \delta_{yy^*} y_t^* + -\gamma_i i_{t-1} + \gamma_\pi \pi_{t-1} + \varepsilon_t^{y} \tag{8}$$

$$y_t^* = \delta_{y^*y} y_t + -\theta_i i_{t-1} + \theta_\pi \pi_{t-1} + \varepsilon_t^{y^*} \tag{9}$$

$$i_t = \phi_\pi \pi_t + \phi_y y_t + \phi_{y^*} y_t^* + \varepsilon_t^{i} \tag{10}$$

and all parameters are assumed to be positive ($\gamma_i, \alpha_y, \phi_\pi, \phi_y > 0$). New equation (9) for y_t^* emerges merely from a mechanical relaxation of assumptions, not theory, and differs from the usual LRAS model based on a production function. Consequently, with no *ex ante* theory to guide the specification, we begin with an equation analogous to the IS curve (8) for y_t. In matrix form, the 4-variable unrestricted SVAR model (SVAR4u) is:

$$\begin{bmatrix} 1 & -\alpha_y & -\alpha_{y^*} & 0 \\ -\gamma_\pi & 1 & -\delta_{yy^*} & \gamma_i \\ -\theta_\pi & -\delta_{y^*y} & 1 & \theta_i \\ -\phi_\pi & -\phi_y & -\phi_{y^*} & 1 \end{bmatrix} \begin{bmatrix} \pi_t \\ y_t \\ y_t^* \\ i_t \end{bmatrix} = B_{4,1} z_{t-1} + \varepsilon_t \tag{11}$$

Equation (11) is not identified because B_0 contains far too many structural parameters (11 versus a maximum of 6). Thus, equation (11) requires additional restrictions or expanded theory and more equations.

To obtain identification of equation (11), we impose the following theoretical restrictions to reduce the number of B_0 parameters to estimate:

- **Real rate**—re-imposing the real rate restriction ($\gamma_{yi} = -\gamma_{y\pi}$ and $\gamma_{yi} = -\gamma_{y\pi}$) eliminates two parameters. Because we show (later) that the data reject the real rate restriction in B_0, imposing $\gamma_{yr} = \theta_{yr} = 0$ eliminates two more parameters.

- **Potential output**—As specified, equation (9) is a linear combination of (8), so B_0 is not full rank. Setting one of the parameters ($\delta_{y^*y}, \theta_i, \theta_\pi$) in equation (9) equal to zero eliminates the singularity but there is little theoretical justification for any individual one. Setting them all equal to zero in B_0 (but not in other lag coefficients) is consistent with an intuitive one-period delayed response of y_t^* to other shocks and eliminates three more parameters.[6]

- **Actual output**—Contemporaneously, it seems plausile that actual output could respond to shocks in potential output fully ($\delta_{yy^*} = -1$), not at all ($\delta_{yy^*} = 0$), or partially ($0 \geq \delta_{yy^*} \geq -1$). Without *a priori* theory, it seems sensible to "let the data decide."

Imposing these restrictions yields an unrestricted 4-variable SVAR (SVAR4u) with impact matrix:

$$B_{4,0}^u = \begin{bmatrix} 1 & -\alpha_y & -\alpha_{y^*} & 0 \\ 0 & 1 & -\delta_{yy^*} & 0 \\ 0 & 0 & 1 & 0 \\ -\phi_\pi & -\phi_y & -\phi_{y^*} & 1 \end{bmatrix}$$

that can be estimated using standard methods. The restricted 4-variable SVAR (SVAR4r)

[6]This restriction is equivalent to ordering y^* first in a reduced-form VAR. In the textbook AD-AS model, long-run aggregate supply (LRAS), y_t^*, is a function of technology and factors of production, but these variables are not in SVAR4u.

with impact matrix

$$B_{4,0}^{r} = \begin{bmatrix} 1 & -\alpha & -\alpha & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & -\widehat{\phi} & -\widehat{\phi} & 1 \end{bmatrix}$$

is obtained by imposing output gap restrictions as follows:

- $\alpha_y = -\alpha_{y^*}$. Whether y_t and y_t^* have different quantitative effects on inflation is a reasonable empirical question. Increases in y_t^* (LRAS) likely reduce aggregate price and inflation, and increases in y_t (AD) likely increase them, but it is unclear theoretically whether these effects are equal with opposite signs.

- $\delta_{yy^*} = 1$. This restriction ensures that any variation in the output gap in the short run is due to movement in actual output. Shocks to potential immediately pass through to actual output, so the output gap does not vary from these potential output shocks.

- $\phi_y = -\phi_{y^*}$. Whether y_t and y_t^* have different quantitative effects on monetary policy is a much less obvious empirical question, but potentially worthwhile asking.[7]

When estimating the SVAR4r, we impose the output gap restrictions in both the impact and lag matrices. When estimating the SVAR4u, we relaxing the output gap restriction in both the impact and lag matrices.

3.3 Future Extensions

The VAR and SVAR macro modeling frameworks are intentionally basic so that econometric veracity can be demonstrated in the data with the simplest, least theoretically constrained framework. However, at least three model extensions are of interest and potentially important. The following subsections briefly discuss the ideas for each future extension.

[7] One reason monetary policy may not react differently to changes in y_t^* is that it's very difficult to discern changes in potential output or productivity growth at high frequencies. (Orphanides and Norden (2002), Orphanides (2003), Arnold (2009)). That said, a leading interpretation of Greenspan's "dovish" monetary policy during the late 1990s is precisely this phenomenon (Gorodnichenko & Shapiro, 2007).

3.3.1 Disaggregated Potential Output

The CBO's complete model of potential output is

$$Y_t^* = A_t^* f(L_t^*, K_t^*). \tag{12}$$

This production function structure can be introduced to the VAR and SVAR frameworks to investigate the dynamics of the components of y_t^*. A key question is, which element(s) of y_t^* is the most likely channel(s) through which LRAS shocks affect the economy? Relatedly, which element(s) were most likely at work during the COVID recession?

The CBO releases data on trend productivity (A_t^*) and potential labor supply (L_t^*) but not potential capital (L_t^*), which is implicitly included in A_t^*.[8]

3.3.2 Fiscal Policy

To be completed...

3.3.3 New Keynesian RE Model

Although they are roughly consistent with a benchmark small-scale New Keynesian (NK) macro model, the VAR and SVAR models exclude expectations of future variables. Estimating the rational expectations version of the NK macro model is an important robustness check; see Galí (2015) for the full model derivation. It also may yield additional insights into how agents' expectations of future actual and potential output differ and whether that heterogeneity matters for the full macro model.

[8]From Shackleton (2018): "Although capital is used in all sectors of the economy, the agency ignores the contribution of capital in its estimates of potential output for the farm and nonprofit sectors. In nearly all cases, the agency's estimate of the potential flow of services from any particular type of capital is directly related to the size of the actual stock; no adjustment is made for business-cycle effects. Consequently, changes in output that result from changes in the intensity of use of capital over the business cycle appear mainly as changes in nonfarm business TFP" (p.21).

4 Econometric Framework

This section describes the data and sample period, econometric model specifications, and methods of estimation and hypothesis testing.

4.1 Data and Sample Period

In our estimation, we use quarterly U.S. Real GDP, the CBO's estimates of U.S. Real Potential GDP, PCE inflation, and the fed funds rate. Our data runs from 1960-2008. We truncate our sample at the end of 2008 for two reasons. First, estimates of GDP and potential GDP are frequently revised for several years after publication. Thus, to ensure we are using data which will not be later revised, we exclude more recent years of data. Second, the Federal Reserve lowered interest rates to the zero-lower bound and began quantitative easing in December 2008. Incorporating this nonlinearity and associated unconventional policy into a VAR requires a different estimation method, as noted in (Mavroeidis, 2021). As these policies are not central to our research question, we exclude them from our sample.

4.2 Model Specifications

In the standard 3-variable VAR framework, the model is stable in its current state: the output gap is stable by construction, and inflation and the nominal interest rate are cointegrated in the long run (Engle & Granger, 1987). However, separating the output gap into actual GDP and potential GDP introduces two non-stationary variables. Further, these variables should be cointegrated *in levels*. To account for this, we estimate our model using a vector error-correction framework using the log-levels of potential and actual GDP, inflation, and the federal funds rate. We use two cointegrating vectors: The first accounts for the cointegration between the level of actual GDP and potential GDP, and a second one to account for the Fisher relationship between inflation and interest rates.

4.3 Estimation and Testing

To evaluate the validity of the output gap restriction, we perform a series of statistical tests. First, we compare the coefficient estimates between the restricted and unrestricted models to identify significant differences. Next, we use Wald tests to assess the validity of individual restrictions, such as the symmetry of coefficients. Finally, we employ log-likelihood ratio tests to evaluate the overall fit of the models. These tests provide a comprehensive framework for assessing whether the output gap restriction is supported by the data.

5 Estimation Results

This section reports the coefficient estimates and results of hypothesis tests for the small-scale macro models.

5.1 Reduced-Form VAR

Table 1 reports the coefficient estimates for the reduced-from VAR4R and VAR4U. Coefficient estimates reveal a consistent significant difference between the restricted and unrestricted model estimates. Specifically, lags of potential output have a larger effect on inflation in the unrestricted model than in the restricted model. Similarly, lags of potential output have a larger effect on actual output in the unrestricted model than in the restricted model. Finally, lags of potential output have a smaller effect on the interest rate than in the restricted model. Each of these differences is statistically significant. As a whole, lags of potential output have a larger effect on inflation, actual output, and potential output than lags of actual output.

Next, in Table 2 we test whether each restriction itself holds via the Wald test. The Wald test fails to reject that inflation responds to lags of the output gap. Similarly, the Wald test fails to reject that the interest rate responds to lags of the output gap, rather than to actual output and potential output uniquely. However, the restriction that changes in potential

output pass through as changes to actual output is statistically significantly rejected. This test, in turn, rejects that short-run variation in the output gap is driven by changes in actual output. In total, the restrictions imposed in B^R do not significantly improve the model's fit, as the combined restrictions are rejected by the log likelihood test.

Table 1: Estimates of $B_{4,1}$ from the restricted and unrestricted VAR models

$\hat{B}^R_{4,1}$				$\hat{B}^U_{4,1}$				$\hat{B}^R - \hat{B}^U$				
.50***	.03	-.03	.04	.52***	.01	-.22***	.02	-.02	.02	.19***	.02	π_{t-1}
(.06)	(.02)	(.02)	(.03)	(.07)	(.03)	(.08)	(.02)					
-.34**	.12	.86***	.11*	-.34***	.20***	1.72***	.19***	0	-.08	-.86***	-.08	y_{t-1}
(.17)	(.07)	(.07)	(.06)	(.17)	(.07)	(.22)	(.07)					
-.01	.02***	.97***	-.01	-.01	.02***	.98***	-.01	0	0	-.1	0	y^*_{t-1}
(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)					
.16	.39***	-.39***	.05	.03	.29***	-.08	.14*	.13	.1	.31***	.09	i_{t-1}
(.20)	(.09)	(.09)	(.07)	(.21)	(.09)	(.27)	(.08)					

Table 2: Tests of the restrictions imposed in the $B^R_{4,1}$ matrix.

Restriction	Wald statistic
$b_{\pi y} = -b_{\pi y^*}$	1.2
$b_{iy} = -b_{iy^*}$	1.5
$(b_{yy} - b_{y^*y}) = -(b_{yy^*} - b_{y^*y^*})$	11.84***
	Chi-Squared statistic
Log likelihood test:	12.09***

5.2 Structural VAR

Table 3 reports coefficients for the impact matrix of the SVAR4R and SVAR4U. Similar to the reduced-form estimation, estimates show a consistent difference between the restricted and unrestricted models on the y^* coefficients. In the unrestricted model, ceteris paribus, inflation increases by 1.04 p.p. contemporaneously following a 1 p.p. shock to potential output, compared to a 0.05 p.p. increase in the restricted model. Similarly, the interest rate increases by 1.78 p.p. contemporaneously from a 1 p.p. shock to potential output, compared to a .38 p.p. increase in the restricted model. Finally, actual output responds more strongly to shocks to potential in the restricted model than in the unrestricted model,

so shifts in potential output can shift the implied output gap contemporaneously. Each of these differences is statistically significant.

Table 3: Estimates of B_0 from the restricted and unrestricted SVAR4 models

$\hat{B}_0^R$				$\hat{B}_0^U$				$\hat{B}_0^R - \hat{B}_0^U$				
1	-.05* (.03)	.05* (.03)	0	1	-.06** (.03)	1.04 (.80)	0	0	-.01	-.99*	0	π_t
0	1	-1	0	0	1	-6.23*** (2.01)	0	0	0	5.23***	0	y_t
0	0	1	0	0	0	1	0	0	0	0	0	y_t^*
-.49** (.21)	-.38*** (.08)	.38*** (.08)	1	-.51** (.21)	-.37*** (.08)	1.78 (.41)	1	-.02	.01	-1.4***	0	i_t

Table 4: Tests of the restrictions imposed in the B_0^R matrix.

Restriction	Wald statistic
$\alpha_y = -\alpha_{y^*}$	9.78***
$\phi_y = -\phi_{y^*}$	1.2

	Chi-Squared statistic
Log likelihood test:	21.50**

The Wald test significantly rejects that inflation responds contemporaneously to the output gap. In turn, actual and potential output have heterogeneous effects on inflation in the short-run. In contrast, the Wald test fails to reject the output gap restriction in the Taylor Rule. Taken as a whole, the restrictions imposed in B_0^R do not significantly improve the fit of the model, so the combined restrictions are rejected by the log likelihood test.

6 Model Dynamics

This section reports the impulse response functions (IRF) for each small-scale macro model.

6.1 Reduced-Form VAR Responses

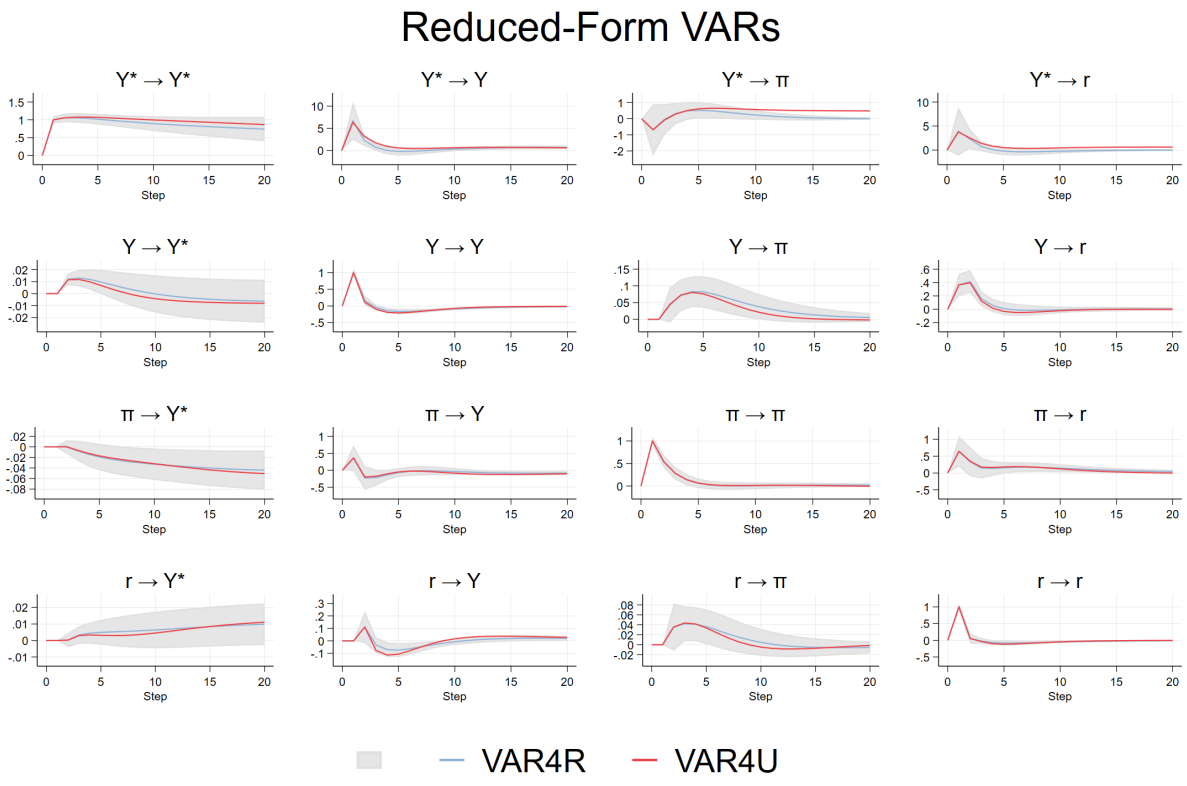
Impulse response functions show more modest differences between the two models. First, shocks to potential output have a persistent effect on potential output, though they raise actual output significantly in the short-run. Potential output shocks lower inflation and raise interest rates in the short-run in both models. However, specifically in the unrestricted model, both inflation and the interest rate settle into a new, higher steady state following a shock to potential output. Both variables are 0.5 percentage points higher in steady state following a shock to potential output. Shocks to potential output also have noticeably different magnitude effects than shocks to actual output. A 1 p.p. shock to potential output shifts inflation by 1 p.p. in the short-run, while a shock to actual output shifts inflation by 0.1 p.p. Similarly, a 1 p.p. shock to potential output shifts interest rates by 4 p.p. in the short-run, while a shock to actual output shifts interest rates by 0.4 p.p. Finally, a 1 p.p. positive shock to potential output actually *increases* the imputed output gap by 5 p.p., while a 1 p.p. positive shock to actual output increases the output gap by 1 p.p.

6.2 Structural VAR Responses

Impulse response functions show more substantial differences between the SVAR4R and SVAR4U models. In the SVAR4U model, actual GDP increases more following a shock to potential output, inflation declines more sharply in the short run before settling into a higher steady state, and interest rates rise more significantly in the short run. The models also show differing dynamics to other shocks as well. Following a shock to actual output, inflation stays elevated for 8 additional quarters in the unrestricted model, while the interest rate stays elevated for 2 additional quarters. The structural VAR models also show notable magnitude differences between shocks to potential output and shocks to actual output. For example, inflation responds by 0.5 percentage points to a one-percentage-point shock to potential output, compared to just 0.1 percentage points for a similarly sized shock to actual output. Additionally, the interest rate changes by four percentage points following a one-

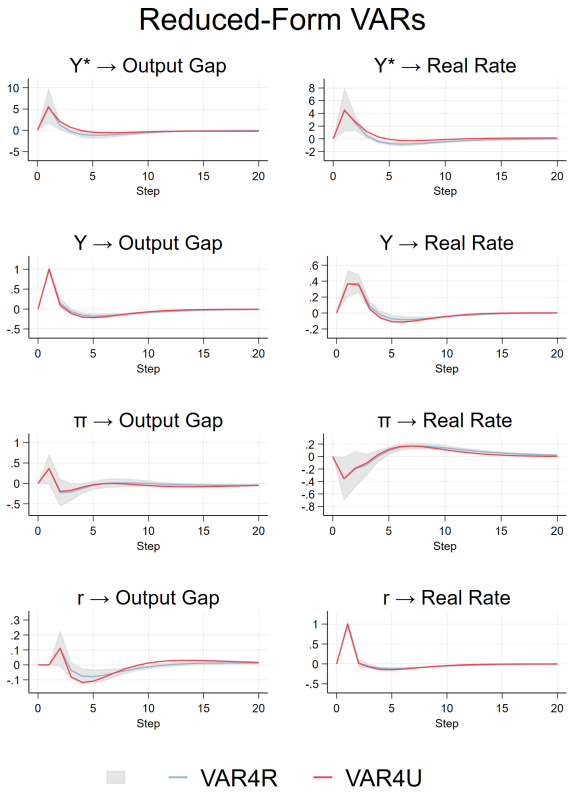
percentage-point shock to potential output, while it changes by only 0.4 percentage points for a shock to actual output. Notably, potential GDP itself is largely unaffected by shocks to other variables in both the short and long run.

Figure 2: Impulse response functions from the 4-variable VAR



There are also notable differences between the restricted and unrestricted models' effects on the imputed output gap and real interest rate. Following a shock to potential output, the output gap *increases* by 5 p.p. while the output gap does not shift in the restricted model. Further, the imputed output gap only increases by 1 p.p. following a shock to potential output. finally, the real interest rate increases more in response to both shocks in the unrestricted model.

Figure 3: Impulse response functions from the 4-variable VAR



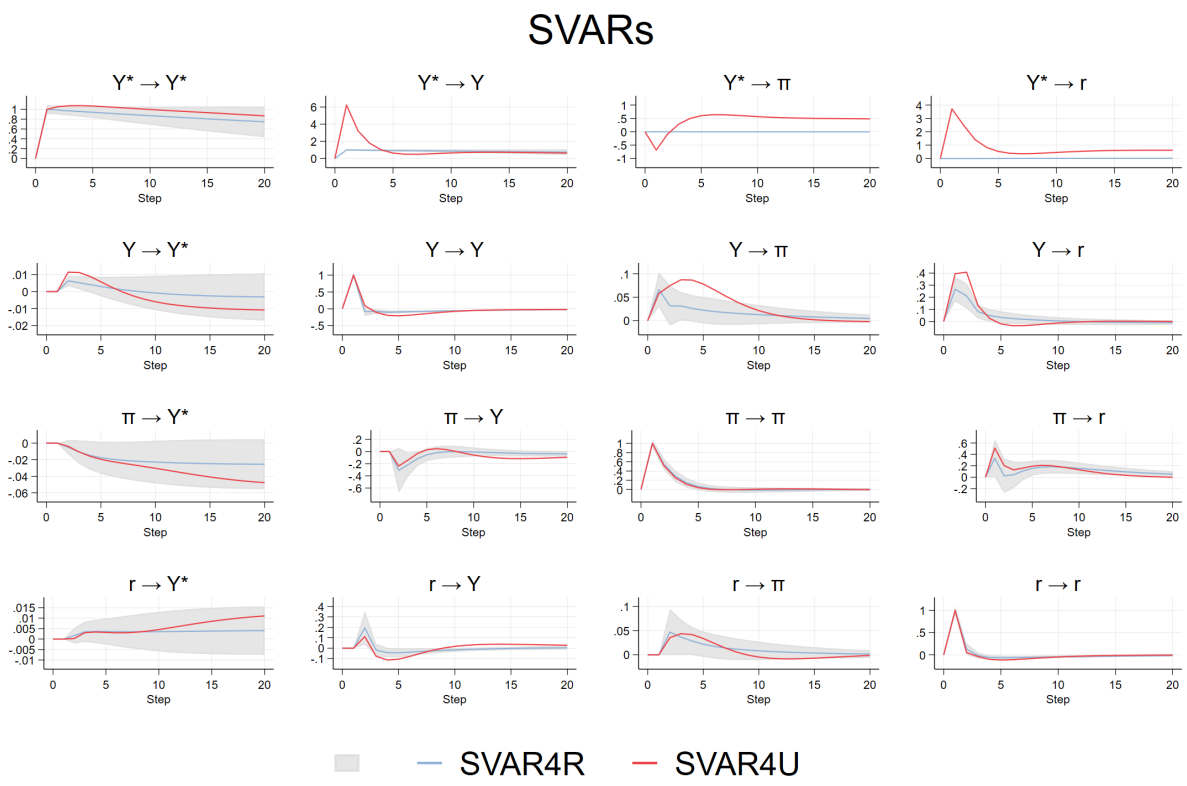
7 COVID-19 Recession Analysis

This section uses the unrestricted SVAR model (SVAR4u) to analyze the COVID recession. Using a basic AD/AS framework, it simulates the estimated SVAR4u model out of sample (2020-2024) for AD versus LRAS shocks. The impulse responses are evaluated to interpret the model's success in fitting the actual data and the Federal Reserve's FOMC forecasts.

7.1 Context and Motivation

The impulse response simulations in the next subsection are motivated by the textbook exposition of the AD/AS framework illustrated in Figure 6. A negative demand shock shifts the AD curve to the left (left panel). In the short run, actual output falls below potential and the aggregate price *declines*, so the shock is *deflationary*. A negative supply shock shifts

Figure 4: Impulse response functions from the 4-variable VAR

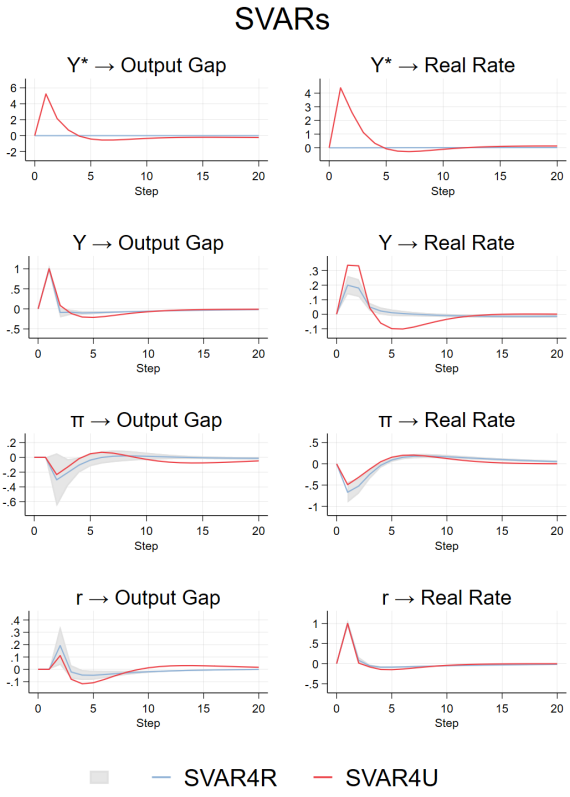


the LRAS curve to the left (right panel). In the short run, actual output also falls below potential but the aggregate price *rises*, so the shock is *inflationary*.

7.2 Model Simulations

As a first step toward understanding the potential role of shocks to LRAS in business cycles, we plot the IRFs from the reduced form VAR for a data-consistent shock to potential output in 2019Q4 and 2020:Q1. As shown in the previous section, small shocks to potential output that are consistent with history and extreme persistence in potential over the business cycle produce very modest responses in model variables. In contrast, the COVID-19 recession involved a drop in output of 1.4% in 2020Q1 and 7.9% in 2020Q2 followed the next quarter by a rebound of 7.8%, leaving output (the output gap) 2.45 percentage points below where

Figure 5: Impulse response functions from the 4-variable VAR

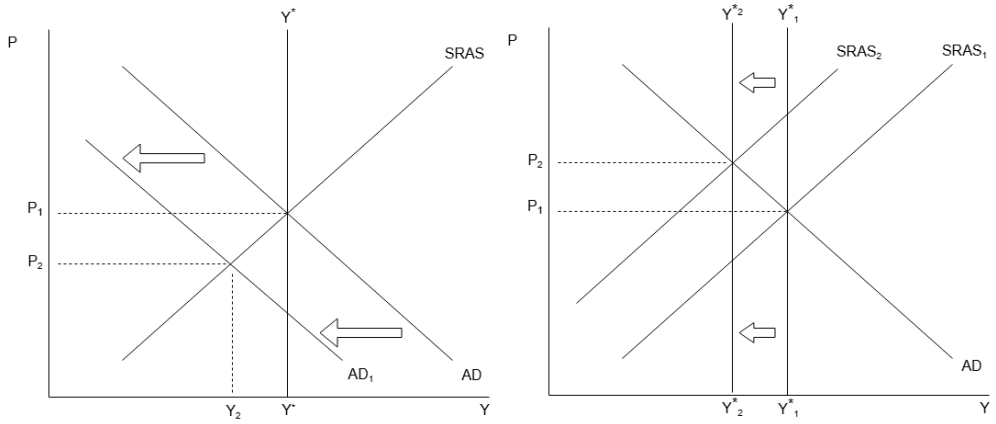


they were prior to the recession.

We simulate the estimated (1960-2007) model with competing AD and LRAS shocks for the period 2020:Q1 through 2023:Q4. Rather than shocking the model with a typical one-standard deviation or 1-percentage point shock, we simulate the COVID shock with data-consistent errors to match 2020:Q1 through and 2020:Q3 exactly. For the LRAS shock, we use the following shock vector:

$$\varepsilon_t^{y^*} = [-1.02,\ -4.00,\ +5.01,\ 0,\ \ldots]$$

with the elements in percent. This shock pattern produces a path for *actual* output, y_t, that matches the data exactly. This pro-factual short-run restriction is necessary to overcome the extreme historical (pre-2008) persistence in model dynamics between y_t and y_t^*, which clearly could not conform to the actual COVID recession developments. For the AD shock,



(a) AD Shock (b) LRAS Shock

Figure 6: Textbook AS/AD interpretations of the Covid recession

we use the actual GDP data. Inflation and interest rate shocks are zero for all periods in both simulations. The SVAR4u model's IRFs govern the responses of all variables.

Figure 7 plots the IRFs for the simulated COVID-19 shocks (blue lines) along with actual data (black dashed lines) and Federal Reserve FOMC forecasts in early 2020 (green dashed lines). The left column depicts the LRAS shock, and the right column the AD shock. Given the pro-factual shock path in early 2020, the SVAR4u model replicates the qualitative behavior of actual GDP reasonably well overall, but economically meaningful differences arise in the other variables.

A closer look at the details is informative. Both LRAS and AD shocks match actual GDP (y_t) quite well, which is surprising at least *a priori*, and both y_t responses are much higher than what the Federal Reserve projected in early 2020. However, the LRAS shock produces impulse responses for inflation and interest rates that are closer to the actual data than does the AD shock. The LRAS shock produces an initially positive inflation response and ultimately higher inflation path than the AD shock, although both responses fail to match the magnitude or persistence of the large inflation runup in 2021-2023. The LRAS shock also predicts much tighter monetary policy, higher than the actual interest rate path until mid-2022 when the Fed tightened more and faster than it expected in 2020.

Although the out-of-sample dynamic simulations of the SVAR4u model reveal interesting

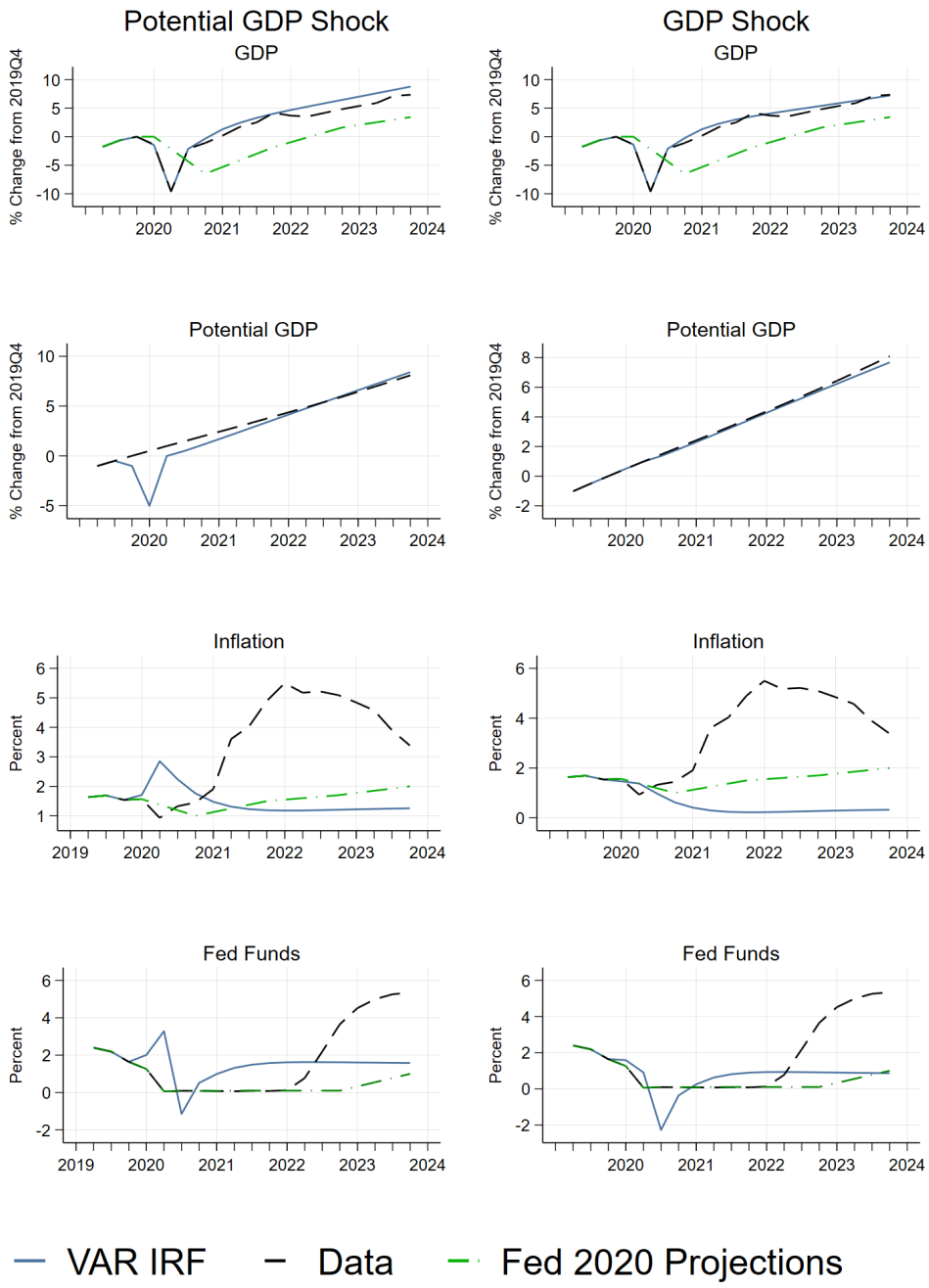


Figure 7: Response functions from COVID-19 simulation. The blue line is the response of the VAR to a large negative shock to Potential GDP in beginning in 2019Q4 and a large positive shock to Potential GDP in 2020Q2. The green line is the Fed's forecast from the June 2020 meeting. The black dashed line is actual data.

distinct difference between LRAS and AD shock paths favoring the former, neither shock matches the full data after COVID. One key reason for this limitation is that the model does not include the massive fiscal stimulus that caused income and consumption to diverge unlike any previous recession. The model needs to include money, which grew more than 40 percent in less than two years, or government bonds, which were monetized in part Fed policy. Another limitation of the model is that it excludes the decomposition of real potential GDP into its key long run components—productivity, labor, and capital—that would facilitate a more detailed interpretation of the LRAS shock. These extensions are left for future research.

8 Conclusion

This study provides evidence that the output gap restriction has not held at business cycle frequencies for most of the post-WWII period. Potential output exhibits short-term fluctuations that have significant implications for macroeconomic dynamics. Inflation and interest rates respond more strongly to shocks in potential output than to shocks in actual GDP, highlighting the distinct roles of these variables in shaping economic outcomes. These findings suggest LRAS shocks are more important than previously thought. This finding challenges traditional views of the output gap and covariance between business cycles and growth, which motivates a reassessment of macroeconomic models and policy frameworks.

The results in this paper have important implications for how to interpret the unusual COVID-19 recession and for monetary and fiscal policies. Policymakers face even greater need to distinguish between shocks to actual and potential output, as these shocks elicit different responses in inflation and interest rates. Failure to account for the independent effects of potential output may result in suboptimal policy decisions, particularly during periods of severe or unusual economic disruption. Future research should build on these findings by incorporating fiscal policy more directly into macroeconomic models, with a focus on understanding the interactions between aggregate demand and supply-side dynamics.

References

Arnold, R. W. (2009). The challenges of estimating potential output in real time. *Federal Reserve Bank of St. Louis Review*, *91*(July/August 2009).

Baqaee, D., & Farhi, E. (2022). Supply and demand in disaggregated keynesian economies with an application to the covid-19 crisis. *American Economic Review*, *112*(5), 1397–1436.

Benati, L. (2007). Drift and breaks in labor productivity. *Journal of Economic Dynamics and Control*, *31*(8), 2847–2877.

Boschen, J. F., & Mills, L. O. (1990). Monetary policy with a new view of potential gnp. *Business Review*, (Jul), 3–10.

Campbell, J. Y., & Mankiw, N. G. (1987). Are output fluctuations transitory? *The Quarterly Journal of Economics*, *102*(4), 857–880.

Campbell, J. Y., & Mankiw, N. G. (2024). Six beliefs i have about inflation. *Journal of Monetary Economics*, *148*(Supplement), 103631.

Campbell, J. Y., & Mankiw, N. G. (2025). A look at inflation in recent years through the lens of a macroeconomic model. https://www.stlouisfed.org/on-the-economy/2025/jan/look-inflation-recent-years-lens-macroeconomic-model

Clark, P. K. (1983). *Okun's law and potential gnp*. Board of Governors of the Federal Reserve System.

Clark, P. K. (1987). The cyclical component of us economic activity. *The Quarterly Journal of Economics*, *102*(4), 797–814.

Coibion, O., Gorodnichenko, Y., & Ulate, M. (2017). *The dyclical sensitivity in estimates of potential output* (tech. rep.). National Bureau of Economic Research.

Crump, R. K., Eusepi, S., Giannoni, M., & Şahin, A. (2024). The unemployment–inflation trade-off revisited: The phillips curve in covid times. *Journal of Monetary Economics*, 103580.

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 251–276.

Fernald, J. G. (2015). Productivity and potential output before, during, and after the great recession. *NBER Macroeconomics Annual*, *29*(1), 1–51.

Fernald, J. G., & Li, H. (2022). The impact of covid on productivity and potential output.

Ferroni, F., Fisher, J. D., & Melosi, L. (2024). Unusual shocks in our usual models. *Journal of Monetary Economics*, 103598.

Galí, J. (2015). *Monetary policy, inflation, and the business cycle: An introduction to the new keynesian framework and its applications*. Princeton University Press.

Gorodnichenko, Y., & Shapiro, M. D. (2007). Monetary policy when potential output is uncertain: Understanding the growth gamble of the 1990s. *Journal of Monetary Economics*, *54*(4), 1132–1162.

Guerrieri, V., Lorenzoni, G., Straub, L., & Werning, I. (2022). Macroeconomic implications of covid-19: Can negative supply shocks cause demand shortages? *American Economic Review*, *112*(5), 1437–1474.

Hall, R. E., & Kudlyak, M. (2023). *The active role of the natural rate of unemployment during cyclical recoveries* (tech. rep.). National Bureau of Economic Research.

Kahn, J. A., & Rich, R. W. (2007). Tracking the new economy: Using growth theory to detect changes in trend productivity. *Journal of Monetary Economics*, *54*(6), 1670–1701.

Kuttner, K. N. (1994). Estimating potential output as a latent variable. *Journal of business & economic statistics*, *12*(3), 361–368.

Martin, R., Munyan, T., & Wilson, B. A. (2015). Potential output and recessions: Are we fooling ourselves? *FRB International Finance Discussion Paper*, (1145).

Mavroeidis, S. (2021). Identification at the zero lower bound. *Econometrica*, *89*(6), 2855–2885.

Okun, A. M. (1970). The political economy of prosperity. *(No Title)*.

Orphanides, A. (2003). The quest for prosperity without inflation. *Journal of monetary Economics*, *50*(3), 633–663.

Orphanides, A., & Norden, S. v. (2002). The unreliability of output-gap estimates in real time. *Review of economics and statistics*, *84*(4), 569–583.

Perry, G. L. (1978). *Potential output: Recent issues and present trends*. Brookings Institution.

Reis, R. (2022). *The burst of high inflation in 2021-22: How and why did we get here?* Bank for International Settlements, Monetary; Economic Department.

Shackleton, R. (2018). *Estimating and projecting potential output using cbo's forecasting growth model*. Congressional Budget Office.