COMMENT BY

VALERIE A. RAMEY This paper by Olivier Coibion, Yuriy Gorodnichenko, and Mauricio Ulate presents surprising new results showing that the leading real-time estimates of potential GDP for the United States and other industrialized countries react to temporary demand shocks. Potential GDP is intended to be an estimate of the maximum sustainable level of output that does not generate inflationary pressure. Because it is a supply-side concept, potential output should not react to demand shocks with temporary effects but should react fully to supply shocks with permanent effects. Coibion and colleagues present convincing evidence that none of the leading estimates of potential GDP satisfies this dichotomy.

Coibion and colleagues have three goals for their paper. Their first goal is to demonstrate that estimates of potential GDP by the various governmental and nongovernmental institutions in the U.S. and other industrialized countries overreact to shocks that have temporary effects on actual GDP and underreact to shocks that have permanent effects on actual GDP. The authors carefully construct real-time databases and use a variety of methods for estimating shocks to show convincingly that leading institutions, such as the Congressional Budget Office (CBO), revise their estimates of potential GDP in response to shocks that are easily identified, even in real time, as temporary. The authors estimate a variety of standard demand shocks, such as monetary and fiscal shocks, first showing that the impulse responses of actual GDP imply temporary effects and then showing that the estimates of potential GDP are revised in response to those shocks. They then estimate supply shocks, showing first that they have permanent effects on actual GDP and then that estimates of potential GDP are not revised sufficiently in response. Achieving this first goal constitutes two-thirds of the paper, and is its heart. These sections of the paper make a substantial contribution: the demonstration is very convincing, and the results are important because estimates of potential GDP are central to numerous quantitative models and are also important guides for policymakers. Perhaps one of the most surprising details in their findings is that estimates of potential GDP by the Federal Reserve's army of Ph.D. economists are virtually indistinguishable from a simple Hodrick–Prescott filter trend and that the Federal Reserve's own estimates of potential GDP are revised based on estimated monetary policy shocks. That is, the Federal Reserve's estimates of potential GDP behave as if monetary policy shocks have permanent supply-side effects, even though the impulse responses of actual GDP show no permanent effects of monetary policy shocks.

The paper's second goal is to explore alternative methods for estimating potential GDP that overcome the problems highlighted in the authors' demonstration. Their main suggested alternative is Olivier Blanchard and Danny Quah's (1989) decomposition of GDP shocks into demand and supply shocks using long-run restrictions, known as the BQ method. Coibion and colleagues show that this measure of potential GDP does not suffer from the same weaknesses as standard measures documented in the earlier sections of the paper. In addition, they explore a variety of other methods based on economic-theory with either alternative long-run restrictions based on theory or Phillips curves.

Finally, Coibion and colleagues' third goal is the production of an alternative measure of the current output gap. Using their implementation of the BQ method, they offer an alternative estimate of current potential GDP and conclude that actual GDP was still more than 5 percent below potential GDP in 2017.

I believe that Coibion and colleagues are very successful in achieving their first goal. Their careful demonstration of the weaknesses of current methods makes it clear that estimates of potential GDP can be improved. Regarding their second goal, their explorations of alternative methods are very promising. I believe that their choice of alternatives is very good. However, as I make clear below, there are remaining challenges with the implementation of their preferred alternative, so more work needs to be done. I demonstrate that key assumptions in their implementation lead to their implausible conclusion that current GDP is significantly below potential GDP. As a result, I do not think their estimates are ready for use by policymakers.

THE PAPER'S ALTERNATIVE METHODS FOR ESTIMATING POTENTIAL GDP To address the weaknesses of the standard estimates of potential GDP, Coibion and colleagues explore alternative methods for estimating potential GDP that can distinguish between shocks that have temporary versus permanent effects on actual output. The main alternative method they explore is the BQ decomposition method. This method uses a bivariate time series model with real GDP and the unemployment rate, and it identifies supply shocks as those shocks that have long-run effects on GDP and demand shocks as all other shocks that have temporary effects. Even if one does not agree with BQ's supply shock–versus–demand shock dichotomy, their method is still useful for separating out temporary from permanent shocks to GDP, which is the key to improving estimates of potential GDP.

Coibion and colleagues also explore other alternatives. For example, they use Jordi Galí's (1999) long-run restriction to identify permanent

shocks to technology; John Cochrane's (1994) permanent income hypothesis-motivated method for using the behavior of consumption to identify permanent shocks to GDP; and a Phillips curve model to infer potential GDP from inflation dynamics. The authors' implementation of all these methods implies much larger current output gaps—that is, actual GDP is farther below potential GDP than those implied by the CBO's estimates and others.

I focus on Coibion and colleagues' implementation of the BQ method because that is their favored method, and that method actually gives a more conservative estimate of the gap relative to their other alternatives. Nevertheless, the authors' particular implementation of the BQ method implies a large gap. Their estimate of potential GDP leads them to conclude that "the gap between potential and actual output in the U.S. increased by about 5 log percentage points between 2007:Q1 (when the gap was likely close to zero) and 2017:Q1, leaving ample room for policymakers to close this gap through demand-side policies if they chose to do so." Thus, their estimates can be seen as an encouragement for policymakers to undertake more demand-side stimuli, even when the unemployment rate is below 4 percent.

IMPLICATIONS OF THE AUTHORS' POTENTIAL GDP ESTIMATES I now demonstrate that Coibion and colleagues' alternative estimates of potential, while avoiding the weaknesses they highlighted for the standard estimates, have a number of implications ranging from questionable to implausible. I argue, however, that the problem is that their estimates are based on questionable auxiliary identifying assumptions that are relatively easy to fix.

Implication 1: Coibion and colleagues' estimates of potential GDP decline as much as the CBO's estimates after the Great Recession. One of Coibion and colleagues' main critiques of the CBO revisions of potential GDP is that they lowered them too much from 2007 to 2017, in response to cyclical fluctuations. Figure 1 of their paper shows how the CBO's estimates of potential GDP changed from 2007 to 2017, and figure 13 shows how their own BQ estimates changed in real time from 2007 and 2017. Consider the revision for the estimate of potential output at the end of their sample, 2016:Q4. Using their data and programs, I calculated that the CBO revised down its estimate of potential GDP in 2016:Q4 by about 0.12 log points, whereas Coibion and colleagues' BQ estimate was revised down by about 0.11 log points over the same period. Thus, both methods lead to the same downward revision in potential GDP. If we believe that Coibion and colleagues' method is accurately capturing only permanent shocks, then their method *validates* the CBO revisions.

Implication 2: The implied natural rate of unemployment is implausibly low. We can combine Coibion and colleagues' estimate of the output gap with Okun's law to calculate the implied natural rate of unemployment. In their paper, Coibion and colleagues conduct this exercise in subsection IV.C. However, they use the older historical estimates of –3 for the parameter on the unemployment gap term rather than the more up-to-date estimates of –2 (Ball, Leigh, and Loungani 2017). Furthermore, they use their estimate of the output gap in 2016:Q4. Because the unemployment rate has fallen so much since then, adding more recent data is instructive.

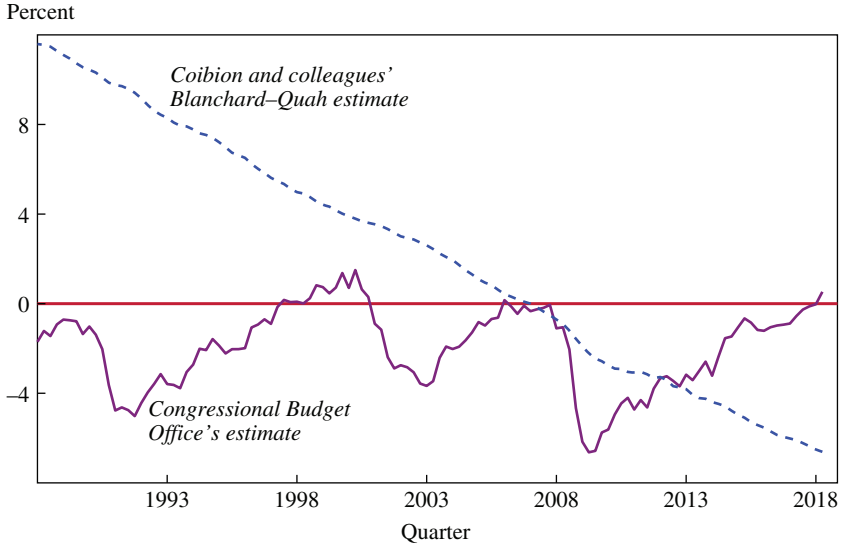
Thus, I update Coibion and colleagues' BQ estimates through 2018:Q2, using their same programs and the same rolling window over the previous 30 years. I find that actual output is about 6.6 percent below their estimate of potential GDP in 2018:Q2. Thus, using their method, I find that actual GDP is farther below potential GDP in 2018:Q2 than it was in 2016:Q4.

The unemployment rate in 2018:Q2 was 3.9 percent. Using Okun's law with a modern unemployment gap coefficient of –2 implies that the natural rate of unemployment in 2018:Q2 was about 0.6 percent. This unemployment rate is below any level ever achieved in the United States, including World War II, and is completely implausible.

Coibion and colleagues argue, however, that the usual Okun's law relationship no longer applies because the employment-to-population ratio in the U.S. fell so much during the Great Recession. Though this is an intriguing possibility, I show below that there is a much simpler explanation for why they estimate such a large output gap and implied low natural rate of unemployment: one of their auxiliary identifying assumptions leads potential GDP to have a significantly higher growth rate than actual GDP in the long run.

Implication 3: Coibion and colleagues' implied output gap has a strong upward trend. As mentioned briefly in discussing the last point, Coibion and colleagues' method for estimating potential GDP implies a bigger output gap in 2018 than at the end of 2016, which seems odd given the fast pace of growth of the U.S. economy and the significant decline in the unemployment rate. This feature led me to inspect Coibion and colleagues' implied output gap for the last 30 years more closely, because they use 30-year rolling regressions to counter possible breaks in trends. In my figure 1, I show the output gap estimated by the CBO and by Coibion and colleagues, where the gap is defined as log actual output minus log potential output so that the gap should be negative at the end of a recession. The CBO's gap behaves as expected, varying cyclically but with no trend. In contrast, the dominant feature of the Coibion and colleagues' implied gap

Figure 1. Estimated Output Gaps, $Y - Y^*$, 1988–2018[a]



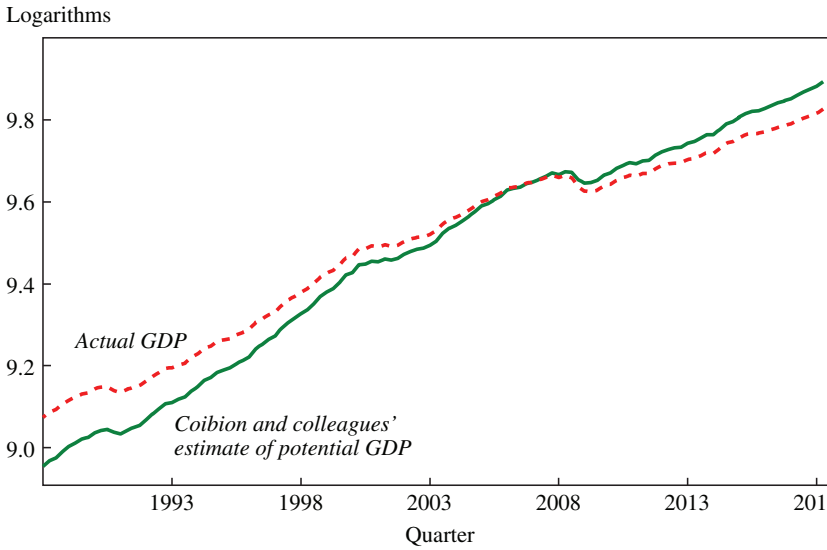
Sources: Author's estimates, using programs from Coibion and colleagues' paper and updated data from the FRED database of the Federal Reserve Bank of Saint Louis.

a. Y = log actual GDP; Y^* = log potential GDP.

is a strong downward trend—the estimated gap declines at a rate of about 0.6 percent per year. According to their estimates, the output gap was very positive in 1988, implying that *actual* output was almost 12 percent above *potential*. However, over time, this gap has narrowed and has become negative. According to the authors' estimates, the output gap is wider now, at about –6.6 percent, than it was at the end of the Great Recession, when it was about –2.2 percent.

This result is a direct consequence of Coibion and colleagues' estimated potential GDP having a much higher trend than actual GDP over the last 30 years. My figure 2 shows the path of both series. Even in the second half of the 1990s, when the growth of total factor productivity surged because of the information technology revolution, they estimate that actual GDP was significantly above potential GDP. The two series cross in 2007, and then the gap becomes negative and widens over time because their estimate of potential GDP grows more quickly than actual GDP. The next section explains which of the assumptions made by Coibion and colleagues lead to this implausible behavior.

Figure 2. Actual GDP versus Potential GDP, as Estimated by Coibion and Colleagues, 1988–2018



Sources: Author's estimates, using programs from Coibion and colleagues' paper and updated data from the FRED database of the Federal Reserve Bank of Saint Louis.

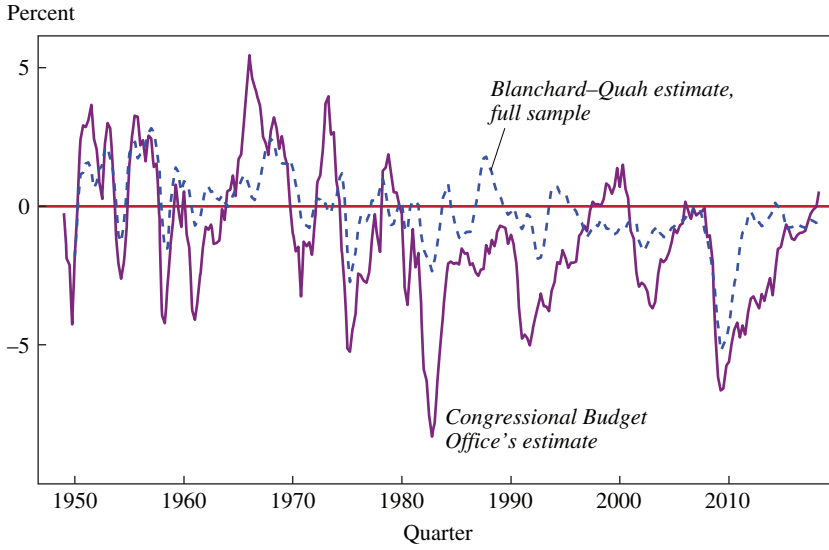
THE BQ METHOD IS NOT ENOUGH TO IDENTIFY POTENTIAL GDP Coibion and colleagues recognize that in order to implement the BQ method to derive a path of potential GDP, they must take a stand on the long-run growth rate of the economy. To see the identification problem, consider the intuitive equation they offer in subsection IV.A of their paper:

$$\Delta \log Y_t = g + \Delta \log Y_t^P + \Delta \log Y_t^C,$$

where $\Delta logY_t$ is the growth rate of actual GDP, g is the long-run growth rate of GDP, $\Delta logY_t^P$ is the growth rate of output due to permanent shocks, and $\Delta logY_t^C$ is the growth rate of output due to temporary shocks. The BQ method assumes that permanent shocks can permanently affect the *level* of GDP, but not the *growth rate* of GDP. Therefore, the BQ method identifies only *deviations* from a long-run path; hence, neither the slope (g) of this path nor the intercept is identified.

Thus, Coibion and colleagues are forced to make two additional assumptions to identify the path. To identify the *slope* of the path, they assume a value of g of 3.1 percent, which equals both the average growth rate of

Figure 3. Estimated Output Gaps, $Y - Y^*$, 1950–2020[a]



Sources: Author's estimates, using programs from Coibion and colleagues' paper and updated data from the FRED database of the Federal Reserve Bank of Saint Louis.

a. The CBO's estimate versus the Blanchard–Quah method's estimate on the full sample.

real GDP from 1977 to 2007 and for the entire post–World War II period. To identify the *intercept* of the path, they assume that potential GDP was equal to actual GDP in 2007:Q1. Also, the CBO's estimated gap is then only about –0.3 percent, so this assumption is close to the CBO's estimates. However, as my figure 2 shows, the slope estimate for g leads the authors' estimate of potential GDP to grow much faster than actual GDP from 1988 to 2018. It is this divergence in growth rates that leads directly to their estimate that output is currently 6.6 percent below potential GDP.

The problem of different growth rates for actual and potential GDP would not occur if g were set equal to the actual growth rate of GDP over the sample used in the estimation. To demonstrate this, I updated the authors' data and reestimated their BQ model back to 1948 and created output gap estimates. These are shown in my figure 3, along with the CBO's estimates. As the figure shows, there is no longer a trend in the gap estimate. However, the two estimates do not move in lockstep. The correlation between the CBO's gap estimate and the BQ gap estimate is about 0.5, suggesting that much could be learned from the differences in the implied gaps.

CONCLUSION Overall, this is an important paper that effectively demonstrates that standard measures of potential GDP overreact to temporary shocks and underreact to permanent shocks. It makes a convincing argument that we can do better, even in real time. The alternative methods explored are promising, but the methods still need work, so any implied gap estimates are "not yet ready for prime time." For now, I think I will stick with the CBO's estimate of the gap, which indicates no slackness in the U.S. economy.

REFERENCES FOR THE RAMEY COMMENT

Ball, Laurence, Daniel Leigh, and Prakash Loungani. 2017. "Okun's Law: Fit at 50?" *Journal of Money, Credit and Banking* 49, no. 7: 1413–41.

Blanchard, Olivier J., and Danny Quah. 1989. "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review* 79, no. 4: 655–73.

Cochrane, John H. 1994. "Permanent and Transitory Components of GNP and Stock Prices." *Quarterly Journal of Economics* 109, no. 1: 241–65.

Galí, Jordi. 1999. "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review* 89, no. 1: 249–71.

GENERAL DISCUSSION

James Stock began by drawing a firmament analogy, wondering if among the stars in the firmament, potential output—*Y**—had any contributions beyond the natural rate of unemployment—*U**. He postulated that in principle, the answer could be yes, because potential output can incorporate capital accumulation, total factor productivity growth, changes in underlying population growth, and changes in the labor force participation rate. This can provide additional information and help explain measures of slackness in the economy and, therefore, thinking about monetary and fiscal policy. However, each of these additional factors has many problems in practice. He acknowledged that perhaps it is plausible to forecast population growth or put aside immigration issues, but there are still ongoing challenges in understanding the labor force participation rate and total factor productivity growth. Although, in principle, it might be possible to get these things right—such as determining the underlying growth rate, and thus making measures of potential output more informative than the natural rate of unemployment—whether this can be pulled off in practice remains doubtful.