Frederic S Mishkin: Estimating potential output


* * *

This conference focuses on measurement issues, and in my remarks I want to focus on one of the most important measurement issues that we at the Federal Reserve and other central banks face: How do we determine whether the economy is operating above or below its maximum sustainable level? That is, how do we estimate the path of potential output?¹

The Federal Reserve operates under a dual mandate to achieve both price stability and maximum sustainable employment. In that context, it is natural to think of potential output as the level of output that is consistent with the maximum sustainable level of employment: That is, it is the level of output at which demand and supply in the aggregate economy are balanced so that, all else being equal, inflation tends to gravitate to its long-run expected value.

The combination of the dual mandate and this definition suggests two reasons that estimating the path of potential output is so central to the conduct of monetary policy. First, to evaluate whether our policies will help achieve the maximum sustainable employment objective of the dual mandate, we need know the level of future output that would be consistent with that objective. Second, the level of output relative to potential output, which is referred to as the output gap, plays an important role in the inflation process. When the actual level of output is above potential output – so that the output gap is positive – labor and product markets are excessively tight; then, if things such as expected inflation and temporary supply factors are held constant, inflation will tend to rise. Conversely, when the output gap is negative and labor and product markets are slack, inflation will tend to fall. Estimates of the future path of potential output are therefore needed to assess whether the projected path of output that is implied by current monetary policy will lead inflation to move in a direction that is consistent with price stability.

Because estimates of potential output are an important part of central bankers' toolkits, the Federal Reserve and other central banks devote considerable resources to getting the best measures of potential output possible. In this talk, I want to explore something that Bismarck warned us we shouldn't want to examine: "what goes into the sausage" – or in this case, what goes into central bankers' thinking about how to estimate potential output.

Broadly speaking, there are three basic approaches to estimating potential output: (1) aggregate approaches; (2) production function, or growth-accounting, approaches; and (3) the newest kid on the block, dynamic stochastic general equilibrium (DSGE) approaches. Let's look at each of these in turn, with the major focus on the production function approach, one to which we at the Federal Reserve currently pay a lot of attention.

(Please note that my comments here reflect my own views and not necessarily those of the Board of Governors or the Federal Reserve System.)

Aggregate approaches

Aggregate approaches to estimating potential output can be thought of as top down approaches because they look at relationships involving aggregate variables and use them to derive measures of potential output. For example, one way of estimating potential output is to assume that if a change in employment or output is sustainable, then it is likely to be permanent. This assumption suggests using univariate statistical methods to identify the permanent component of changes in output, which could then be viewed as a reasonable measure of potential output. Examples of such an approach include the work of Beveridge and Nelson (1981) and Clark (1987). Although univariate approaches to measuring potential output have the advantage of simplicity and can provide a feel for what potential

¹ I want to thank Andrew Figura, Charles Fleischman, John Roberts, and William Wascher for their helpful comments and assistance on this speech.
output might be, they suffer from several disadvantages. First, they require a variety of statistical assumptions about which economic theory provides little guidance— for example, about the correlation between permanent and transitory components or whether the permanent component should be modeled as a random walk. Perfectly sensible alternative assumptions can lead to very different estimates of potential output. Second, these purely statistical approaches do not tell us whether this measure of the permanent component of output movements provides information about the most important aspect of potential output from a central banker’s perspective— namely, its association with a stable rate of inflation.

In measuring potential output, we therefore need to bring in some economics. One potentially valuable economic relationship we can use is the “natural rate” version of the Phillips curve, which followed from the seminal research of Nobel prize winners Milton Friedman (1968) and Edmund (Ned) Phelps (1967). Friedman and Phelps demonstrated that there should not be a long-run tradeoff between inflation and unemployment and that the economy will gravitate to some natural rate of unemployment in the long run no matter what the rate of inflation is. In other words, the long-run Phillips curve is vertical, and attempts to lower unemployment below the natural rate will only result in higher inflation.

According to the natural-rate hypothesis, there is a natural rate of unemployment—which also more commonly referred to as the NAIRU (non-accelerating inflation rate of unemployment) — at which inflation tends to gravitate to its long-run expected value. A natural rate of output— that is, potential output— corresponds to the NAIRU. The difference between actual and potential output, the output gap, tells us whether inflation will tend to move up or down, holding things like inflation expectations, energy prices, import prices, and so forth constant. The natural-rate hypothesis thus suggests that potential output can be estimated from a multivariate approach in which potential output is an unobserved component in the relationship between inflation and the output gap. Kuttner (1994) provides a good example of this approach.

An alternative approach involves deriving the NAIRU directly from estimates of Phillips curves and then using Okun’s law—which relates the output gap to the unemployment gap (the actual unemployment rate minus the NAIRU) — to estimate potential output. These multivariate approaches are reasonably simple and make intuitive sense, but they also have serious drawbacks. First, they require that the specification of the Phillips curve is correct. For example, the model needs to correctly characterize the relationship between the unemployment rate gap and inflation dynamics while taking into account how inflation expectations are formed, and it should not leave out any other variables that have an impact on inflation. Indeed, many economists criticize the Phillips curve, with some even declaring it dead.

Second, using Okun’s law to derive potential output requires an appropriate specification for the dynamics of the relationship between output and unemployment gaps. However, cyclical fluctuations in productivity and labor supply can complicate this relationship. Moreover, Okun’s law can be thrown off course during periods of unanticipated structural change in the economy, such as the early 1970s, when U.S. productivity growth slowed. As a result of these influences, Okun’s law has not

2 The term "NAIRU" comes from a paper by Nobel prize winner Franco Modigliani and Lucas Papademos, now the vice president of the European Central Bank (Modigliani and Papademos, 1978). Although NAIRU and the natural rate of unemployment are frequently used interchangeably, there is a subtle distinction between the two. The natural rate of unemployment is the rate at which inflation would tend to gravitate to its long-run expected value, while NAIRU, as originally defined, is the unemployment rate at which inflation will have no tendency to move up or down. Depending on what shocks are hitting the economy, the NAIRU could deviate from the natural rate of unemployment (for example, see Estrella and Mishkin, 1999). Because the NAIRU terminology is more common than the natural rate of unemployment terminology, I am using the NAIRU terminology even though I think that the natural rate of unemployment terminology is more accurate.

3 Other examples of the multivariate approach include Apel and Jansson (1999), Cochrane (1994), and Dupasquier, Guay and St-Amant (1999).

4 Okun’s law was originally specified as the relationship between real GDP growth and changes in the unemployment rate (see Okun, 1962).

5 See Gordon (1982 and many subsequent papers) for a description of the “triangle” model of inflation. The three sides of the triangle are inflation inertia (captured by the lags of inflation), excess demand (measured by the unemployment rate gap or GDP gap), and supply shocks—such as the relative prices of imports, food, and energy. Following George Perry’s (1970) early work, it is common to adjust the NAIRU for changes in the composition of the labor force. Falllick, Fleischman, and Pingle (2006) estimate that shifts in the demographic composition of the labor force can explain a decline in the unemployment rate of nearly 1 percentage point between 1977 and 2006.

6 See, for example, Atkeson and Ohanian (2001).
always been a reliable guide to the relationship between the unemployment gap and the output gap and thus has not always been the most useful guide for estimating potential output. Finally, even if the Phillips curve and the Okun’s law relationships are specified correctly – a big if – the statistical uncertainty about the estimates of the NAIRU, and therefore also about potential output, is very large – certainly larger than a policymaker would like. For example, the estimates of Staiger, Stock, and Watson (1997a and b) of the 95 percent confidence interval for the NAIRU were as much as 3 percentage points wide. Thus estimates of the NAIRU, in isolation, provide policymakers with little real-time insight for assessing the effect of labor markets on inflation pressures.

Production function (growth-accounting) approaches

Because of the shortcomings of the aggregate approaches described above, some researchers estimate potential output using a production function approach that generates an estimate of potential from the underlying factors of production. This approach is sometimes referred to as “growth accounting” because, after the log of a production function is differentiated, output growth can be expressed as a weighted average of the growth of factor inputs – that is, capital services and labor input (hours worked and labor quality) – and a residual – multifactor (also called total factor) productivity growth. The NAIRU concept still plays an important role in this approach because it helps to determine the level of potential output.

A major advantage of the growth-accounting approach is that it focuses on the various factors that drive growth in potential output, rather than simply on the historical behavior of output growth or on the historical relationship between output and labor inputs as in Okun’s law. The disaggregated nature of the growth-accounting approach means that more data can be used to estimate potential output. These additional data are likely to be particularly valuable when the economy is undergoing major structural changes – for example, the productivity slowdown starting around the early 1970s, the surge and subsequent slowdown in population growth from the baby-boom generation’s entry and (now) exit from the labor force, the remarkable upsurge in labor participation of females in the 1970s and 1980s, and the pickup in productivity growth starting in the second half of the 1990s. Given these advantages, it is not surprising that the growth-accounting framework is widely used to estimate the path of potential output, both by central banks such as the Federal Reserve and by academic researchers.

In its simplest formulation, the growth-accounting framework characterizes output growth as the sum of the growth rate of raw labor hours and the growth rate of output per hour, that is, labor productivity. In turn, the growth rate of labor hours is described as the sum of population growth (civilian non-institutional population aged sixteen and older), the rate of change of the labor force participation rate, the rate of change in the employment rate, and the rate of change in the number of hours in the average workweek. Labor productivity is decomposed into the contributions of capital deepening (the marginal product of capital – typically estimated as capital’s share of income – times the growth rate of capital services per hour), changes in labor quality, and the growth rate of multifactor productivity.

To obtain estimates of potential output growth, researchers can use economic analysis to estimate the individual components above that go into the growth-accounting framework. Much research has been conducted along these lines in the Federal Reserve System and elsewhere. For example, Aaronson and others (2006) examine what variables influence individual decisions to participate in the labor market (birth cohort, age, sex, number and age of children, and so forth). They then use this information together with the relative size of different cohorts to show that the aging of the baby-boom cohort can explain much of the decline in labor force participation since 2000 and why the participation rate is likely to continue to decline further in the future. Given slower projected population growth and what appears to be a downward trend in the number of hours in the average workweek, these results suggest that potential output growth will be slower than it otherwise would have been. Of course, such projections are subject to uncertainty, and economists hold a range of views about the prospects for future labor force growth. One important unknown is whether increases in longevity and better health

---

7 See, for example, Altig, Fitzgerald, and Rupert (1997), Rudebusch (2000), and Groshen and Potter (2003).
8 Because official series for productivity for all sectors of the economy are not available, the growth-accounting framework most often focuses on the private nonfarm business (NFB) sector, which in the United States accounts for more than three-quarters of total output. Productivity for the overall economy is then derived from estimates for the nonfarm business sector and then cruder estimates for the other sectors (government, farm, housing services, and households and institutions) generally based on the univariate and multivariate statistical approaches described earlier.
will boost the labor force participation rates of older individuals more than currently embedded in these projections.

Similarly, economists at the Fed have been at the forefront of research on why labor productivity growth ratcheted upward in the second half of the 1990s and continued at a surprisingly robust pace over the first half of this decade. Oliner, Sichel, and Stiroh (2007), for example, find that the IT (information technology) sector has been a key element in the higher productivity growth that we have been experiencing since the mid-1990s. But they also find that the sources of growth in the second half of the 1990s were quite different than in the period after 2000. In the 1995-2000 period, labor productivity growth was driven by substantial gains in multifactor productivity in the tech sector, which, in turn, led to sharp reductions in the prices of high-tech equipment and stimulated investment in this type of capital in other sectors of the economy. Since 2000, high productivity growth has apparently been driven importantly by industry restructuring in response to pressures on profits (the firms that saw the sharpest drops in profits were those that had the largest gains in labor productivity) and by a reallocation of material and labor inputs across industries. Their estimate for the current trend labor productivity growth is centered around 2-1/4 percent, but they find large uncertainty around this estimate, with a 95 percent confidence interval that ranges from 1-1/4 percent to 3-1/4 percent. Roberts (2001) estimates even larger uncertainty around structural labor productivity growth.

Despite the advantages of the growth-accounting framework, it still presents some difficulties. For one, there is, as I just highlighted, a large degree of uncertainty surrounding the estimates of the components that go into the growth-accounting formulas. In addition, the growth-accounting framework requires a substantial amount of data, some of which are not especially reliable. For example, it is especially difficult to measure the growth rate of capital services, and the Bureau of Labor Statistics releases its initial estimates with a lag of at least one year. In addition, the bureau’s two measures of employment, one derived from a survey of firms and the other from a survey of households, often provide very different pictures of what is happening in the labor market. Also data for capital services, labor composition, and multifactor productivity are not readily available for all sectors of the U.S. economy.

Because it is so difficult to reliably estimate potential output using either the aggregate or the growth-accounting approach, it should come as no surprise that we at the Federal Reserve use a lot of judgment in constructing our estimates of potential output. In particular, we see judgment as playing three important roles in our procedures. First, it enables us to take account of the effects of structural changes in the economy that cannot be modeled directly. Second, it allows us to deal with model misspecifications that cannot be corrected. Third, we can use judgment to correct for measurement errors or inconsistencies in economic data. For example, we judgmentally adjust model-based estimates of the NAIRU to account for movements in the unemployment rate unrelated to changes in labor market slack. Of these, the most important has been the shift in the demographic composition of the labor force, driven largely by the entrance and subsequent maturation of the baby-boom generation. But economists have pointed to a number of other factors that would influence the NAIRU as well.

Another example relates to the way in which we estimate the trend growth rate of multifactor productivity. In particular, estimates of trend multi-factor productivity growth tend to be sensitive to the choice of modeling strategy, a problem that became particularly apparent during the “jobless recovery” of the early 1990s, when the normal relationship between output growth and employment growth appeared to break down. Statistical filtering models like those described by Roberts (2001) tended in real time to attribute much of the weak employment growth to an acceleration in trend productivity. Later on, however, after employment recovered, it became clear that the unusual behavior of employment during that period was explained better as a temporary reluctance by firms to hire than as a step-up in the rate of trend productivity growth. In a similar vein, it may at times make sense to down weight more-recent estimates of the data used in filtering exercises to account for the possibility of future revisions to the data.

Finally, it can often be useful to look at Okun’s law as a check on the estimates of potential output derived from the growth-accounting approach. Although, as I noted above, one should not expect Okun’s law to hold from quarter to quarter, the relationship is relatively robust over longer periods, and

---

9 See also Oliner and Sichel (2000, 2002), Corrado and Sifman (1999), and Jorgenson, Ho, and Stiroh (2004), among others.
a persistent deviation in the unemployment rate from that predicted by Okun's law might call into question the estimated trends in one or more of the components in potential output.

**Dynamic stochastic general equilibrium approaches**

The real business cycle literature, which started with the work of Nobel Prize winners Finn Kydland and Edward Prescott (1982), features optimizing agents and emphasizes the role of technology shocks in explaining both economic growth and business cycles. Dynamic stochastic general equilibrium (DSGE) models contain many features of the earlier real business cycle literature, but, because they allow for rigidities and imperfections in markets, they are often referred to as New Keynesian models. The New Keynesian DSGE models provide more-realistic, yet still theoretically elegant, representations of the economy, and their development has been an exciting area of research in macroeconomics in recent years.

New Keynesian DSGE models provide a somewhat different, but complementary, perspective on the definition of potential output than the one I outlined at the beginning of this speech. In particular, we might think of potential output as the level of output that an economy could attain if the inefficiencies resulting from nominal wage and price rigidities were removed — that is, if wages and prices were fully flexible. The definition of potential output as a flexible price equilibrium has much in common with the more conventional definition I discussed earlier because over time prices (and wages) do gravitate toward their equilibrium levels. As a result, the DSGE definition accords with the idea that potential output is the level of output at which inflation tends neither to rise nor fall.

That said, the DSGE view of potential output also has important differences with the earlier approaches to estimating potential output. Although research on using DSGE models to estimate potential output is in its infancy and so should be read cautiously, papers such as Neiss and Nelson (2005) and Edge, Kiley, and Laforte (2007) are finding that the properties of potential output and output gap fluctuations can be quite different from conventional measures. For example, in many DSGE models, potential output can undergo swings over the business cycle, a result that should not be surprising considering that the early real business cycle models viewed the business cycle as being primarily an efficient response to shocks to the economy. In addition, fiscal policy shocks, changes in households' preferences with regard to saving and consumption, changes in preferences about leisure that affect labor supply, and terms-of-trade shocks can all cause potential output to fluctuate. In contrast, growth-accounting approaches to estimating potential output generally assume that such shocks have no important effects on potential output at business-cycle frequencies. As a consequence, their estimates typically have smaller fluctuations than measures of potential output derived from DSGE models, and thus the output gaps in the current generation of DSGE models tend to be less variable than conventional measures and can be quite different for particular periods.

Although the research on DSGE models is promising, measures of potential output and the output gap from these models are controversial. The DSGE measures of potential output are far more model dependent than more-conventional measures because they depend on the estimated parameters of the model and on the model's estimates of the structural shocks hitting the economy. As a result, DSGE models with different characterizations of the economy's underlying structure can produce substantially different estimates of potential output. This is apparent, for example, in the large differences between the potential output measures in the DSGE models of Neiss and Nelson (2005) from those in Edge, Kiley, and Laforte (2007). Moreover, DSGE models often require strong assumptions to identify the shocks to potential output from model equation residuals. The finding that these models imply smaller and less persistent output gaps than traditional models may simply reflect the fact that inefficiencies other than price rigidities, such as real wage rigidities, are important for output fluctuations. As a result, some policymakers have been quite critical of the implication of DSGE models that a substantial fraction of business-cycle fluctuations are efficient and so do not require a response from monetary policymakers.

---

10 See, for example, Goodfriend and King (1997) or Woodford (2003).
11 For example, Blanchard and Gali (2007) have proposed a model with real wage rigidities that, if incorporated into a DSGE model, would likely show output gap estimates that are more similar to traditional gaps.
12 See, for example, Bean (2005).
Implications for policy

Now that we have looked inside the sausage of estimating potential output, I hope you have not lost your appetite for thinking about what these measurement issues mean for monetary policy. As I indicated earlier, considerable uncertainty surrounds the measures of potential output derived from any of the approaches I have discussed. In addition, there is also what economists call Knightian uncertainty (named after the famous University of Chicago economist Frank Knight) – the fact that we are not even sure of the appropriate modeling approach to measure potential output. Adding even more to the uncertainty of potential output measures are (1) that the observable data do not always correspond to the data we would like to have to produce measures of potential output and (2) that initial estimates of observable data can subsequently be revised substantially, resulting in a very different picture of what is happening to potential output and the output gap. Orphanides (2001) points out that output gaps were grossly mismeasured in the 1970s, in part because the initially published data did not reflect the true state of the economy.\(^{13}\)

Where does the high uncertainty about actual and potential output leave us at central banks? Does it mean that we should abandon our focus on potential output and output gaps in making decisions on monetary policy? I think not. For better or worse, we cannot escape the need for information on output gaps so that we can forecast the future path of inflation and evaluate the current setting of our monetary policy instruments. However, we also need to recognize that because measures of potential output and output gaps are so uncertain, we must always be aware that they might be providing misleading signals as to the future course of inflation and the appropriateness of the stance of policy. In assessing whether there is slack in the economy, we at central banks look not only at our estimates of output gaps but also at a wide range of indicators drawn from the labor, product, and financial markets to provide us with a perspective on the balance of supply and demand in the economy. Most important, the substantial uncertainty in our measures of potential output implies that we need to be cautious about taking on board the implications of our current estimates of the output gap. For example, if inflation is moving in a different direction than the output gap would suggest, then we should take seriously the possibility that our output gap measure is not providing us with reliable information.

The bottom line is that we must never take our eye off of the inflation ball. Good policymaking requires that we acknowledge what we are unsure about, and this requirement applies particularly to measures of potential output.

References


---

\(^{13}\) See also Orphanides and others, (2000) and Orphanides and van Norden (2002).


