Extracting the Signal From the Noise

Tips for Interpreting Macroeconomic Data

BY JASON FURMAN

ILLUSTRATIONS BY RAUL ARRIAS

IN THE EARLY STAGES OF THE GREAT DEPRESSION, policymakers in Washington were faced with a profound gap in their understanding of the state of the U.S. economy: no one actually knew how many Americans were out of work. Aside from some attempts in the decennial census, before the 1930s no government agency had regularly measured the number of individuals seeking work in the United States.

This inability to measure basic economic conditions may seem shocking to observers of the U.S. economy today. The federal statistical agencies – the Bureau of Economic Analysis, the Bureau of Labor Statistics (BLS) and the Census Bureau – and a variety of other public and private entities now provide a wealth of economic data on an annual, quarterly, monthly and sometimes even weekly or daily basis.

Yet, while we no longer must cope with the void that policymakers faced in the 1930s, the mountain of data available creates its own problems. Perhaps chief among them is that we can sometimes ask too much of the data while doing too little to put it in context. The conflicting demands for timely reporting of data and their accuracy and completeness make it necessary to be cautious in interpreting the numbers.

For example, the BLS originally reported that the economy added 38,000 jobs last May, which could have led an observer to believe the economy was slowing markedly since job growth had averaged over 200,000 a month in both 2014 and 2015. But then in June, according to the Bureau's initial estimate, the economy added 287,000 jobs – a boom.

The truth is that, at a monthly frequency, it is difficult to accurately measure the vitals of the economy, and placing much weight on monthly data when they are first released can lead one seriously astray in assessing what's happening.

It is not just that numbers bounce around from month to month; seemingly comparable measures can offer divergent readings even for the same period. The United States measures economic output in two different ways that in theory provide different routes to the same destination: Gross Domestic Product and Gross Domestic Income. In the second quarter of 2015, the economy grew a disappointing 0.6 percent, according to one, and a solid 2.6 percent according to the other. In this case the difference between these two numbers was simply statistical noise, a reminder that these statistics are an imprecise way to measure the economy's temperature at a quarterly frequency.

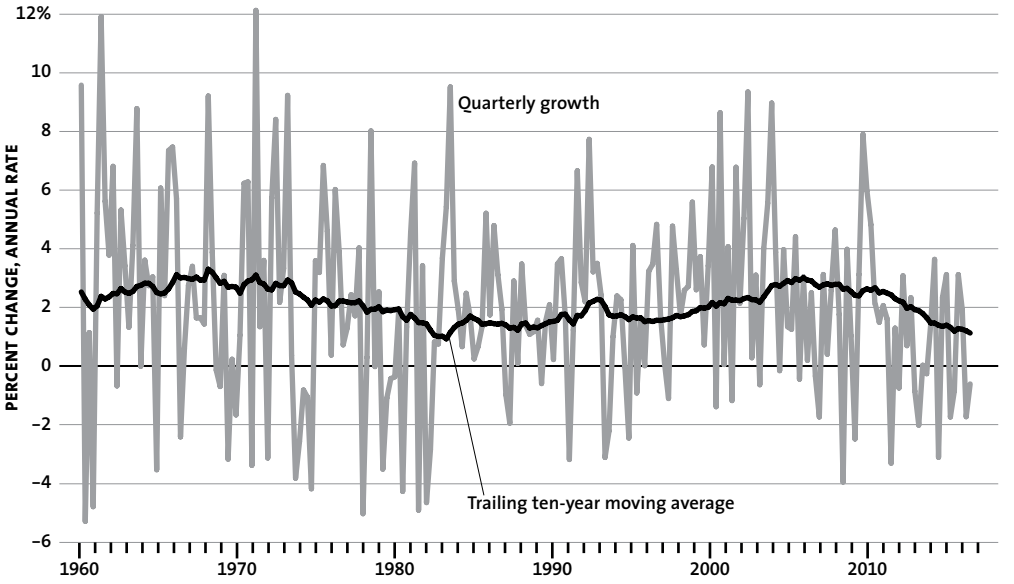
These are not just academic issues. How individuals and institutions (and in some cases, computer algorithms) interpret and react to economic data influences economic policy as well as private consumption and investment decisions.

In the midst of the economic crisis, for example, economic growth for the fourth quarter of 2008 was initially estimated at –3.8 percent and job losses in November 2008 were originally estimated at 598,000. These data points affected perceptions in Washington of what constituted an appropriate fiscal policy response. However, the estimates would later be revised down to much grimmer numbers (–8.2 percent growth and 791,000 jobs lost), which, had they been known earlier, might well have led to a proposal for more stimulus.

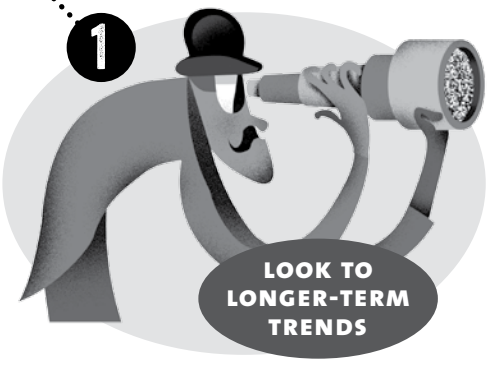
Here, I offer seven lessons to help guide those trying to make sense of the wealth of economic data available today, many of them drawing on analytical work by the Council of Economic Advisers. I also provide some applications of these lessons that have proved most valuable in understanding the economy. But all of this has a simple bottom line: when assessing the overall health of the economy, never read too much into a single data snapshot. Rely, instead, on data series over substantial periods and in the context of what other data suggest is happening.

JASON FURMAN is the chairman of President Obama's Council of Economic Advisers.

FIGURE 1: LABOR PRODUCTIVITY GROWTH, NON-FARM BUSINESS SECTOR



SOURCE: Bureau of Labor Statistics, Productivity and Costs; CEA calculations

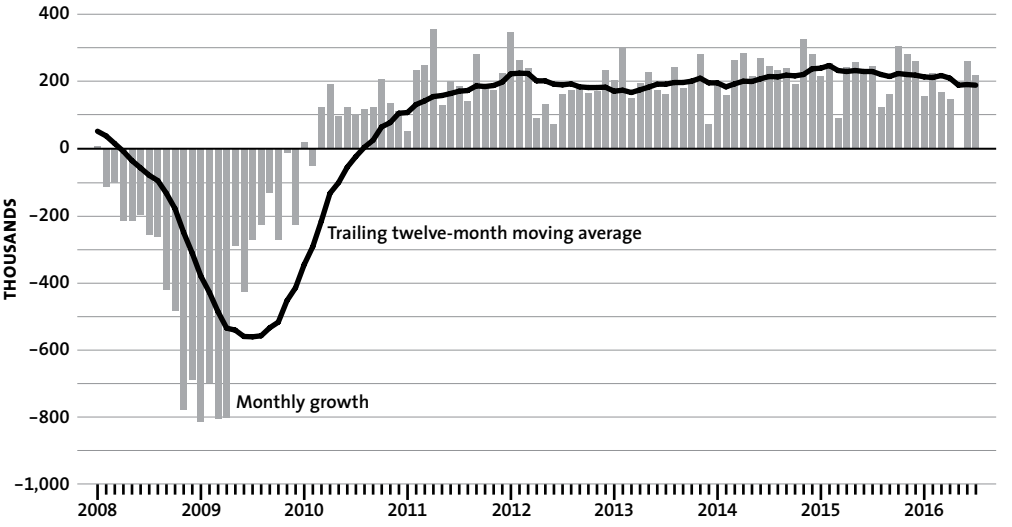


UNDERLYING ECONOMIC REALITY, as well as our attempts to measure it, exhibits substantial volatility. Hence, as important as it can be to gauge turning points in prices, employment, output and the like on a frequent basis, it is also important not to lean heavily on any single data point – or even on a combination of data points – because our measures are nowhere close to perfect. This is not the fault of the statistical agencies, but simply due to the inherent complexity of a vast and rapidly changing economy like that of the United States.

More data over longer periods make it easier to disentangle underlying trends from transitory noise. While there are a variety of sophisticated statistical techniques to smooth economic data, a simple moving average that weights past as well as current numbers equally offers a reasonable way to assess trends.

Take labor productivity, a measure of how much output is produced by an average hour of labor. Measured productivity growth is extremely noisy – that is, full of spurious volatility – at a quarterly frequency, and we largely look to it to answer longer-run questions about the economy. Moreover, there is some evidence that the best predictor of productivity growth is a long-term average of past productivity growth. All of this suggests that, at a minimum, productivity growth should be assessed with something like a trailing 10-year moving average as shown in Figure 1.

FIGURE 2: PRIVATE-SECTOR PAYROLL EMPLOYMENT, 2008–2016



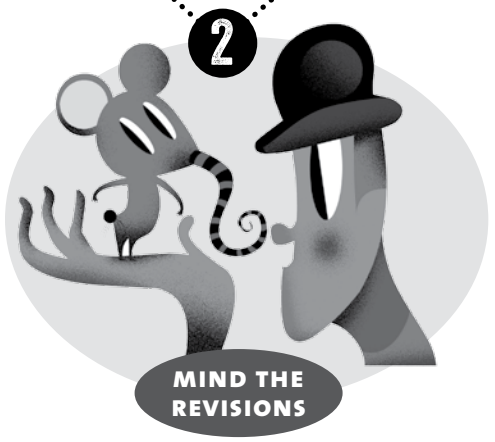
SOURCE: Bureau of Labor Statistics, Current Employment Statistics, CEA calculations

FIGURE 3: WEEKLY UNEMPLOYMENT INSURANCE CLAIMS, 2012–2016

SOURCE: Department of Labor

Payroll job growth is less volatile than productivity growth and thus can be examined in the context of a shorter window like the 12-month trailing moving average shown in Figure 2. From 2012 to the end of 2015, the 12-month moving average of private-sector job growth held steady at about 200,000 per month – a much more accurate picture of the economy than the excessive optimism suggested in the many months when job growth came in above that average or the excessive pessimism of news reports in the many months when job growth fell well short of that average.

Averaging over time is essential with higher-frequency data, like initial claims for unemployment insurance. Initial claims are compiled weekly from administrative data from state offices, so they are not subject to the same measurement error as data derived from sample surveys, such as estimates of job growth. But the series bounces around from week to week, with dramatic movements in both directions that can mislead anyone trying to get a clear picture how many Americans are involuntarily out of work. Last May, for example, initial claims spiked for exactly one week entirely because an unusual law in New York permits many public school employees to claim benefits for their time off during spring break. Using a four-week moving average helps avoid some of the zigzags, as shown in Figure 3, giving a more stable picture of recent trends.



IT'S STANDARD PRACTICE for the statistics agencies to issue revised estimates of economic data that incorporate new information as it becomes available – a fact that is easy to miss, given that these revisions can occur months or even years after the initial reporting. For example, with each month's release of employment data, the Bureau of Labor Statistics also revises the prior two months' estimates of job growth. These revisions are often large and economically meaningful, especially around economic turning points. Estimates of monthly job growth are then revised once a year for the next five years. For example, in September 2011, the Bureau of Labor Statistics reported that job growth in August had been zero, a striking number that fueled concerns the economy was headed into a double-dip recession. But the latest revised estimate for job growth in that month is a far-less-concerning 107,000.

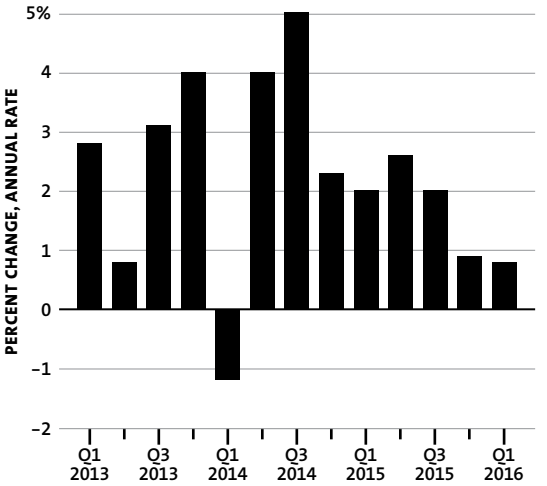
Some of the clearest instances of revisions changing the economic narrative come from the Bureau of Economic Analysis' corrections to quarterly GDP data. When the advance estimate for GDP growth in a given quarter is published, the bureau does not yet have all of the timely data on international trade, business inventories and spending on services; thus, the agency must use projections based on statistical modeling to pencil in more than half of the data. Even nearly three months after the quarter ends, it still has to use trends or indirect indicators to estimate components that comprise about one-third of GDP.

TABLE 1: EXAMPLES OF SUCCESSIVE REVISIONS TO REAL GDP GROWTH (PERCENT CHANGE, SEASONALLY ADJUSTED ANNUAL RATE)

	2001:Q4	2015:Q1
First Estimate	0.2	0.2
Second Estimate	1.4	–0.7
Third Estimate	1.7	–0.2
Estimate as of 8/16	1.1	2.0

SOURCE: Bureau of Economic Analysis, National Income and Product Accounts

FIGURE 4: REAL GDP GROWTH, 2013-2016

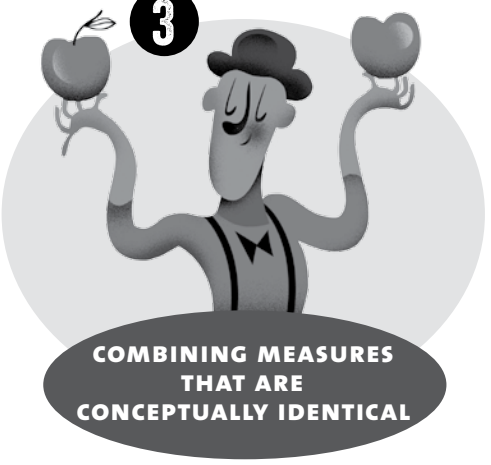


SOURCE: Bureau of Labor Statistics, National Income and Product Accounts

The Bureau of Economic Analysis makes additional revisions to several years' worth of data each July. Together, these revisions can have a dramatic effect on measured economic growth, as shown in Table 1. In the fourth quarter of 2001, the bureau's advance estimate of real GDP growth (calculated as an annual rate) was a tepid 0.2 percent. By the third estimate, this had been revised upward to 1.7 percent; subsequent annual revisions brought it back down to 1.1 percent. In the first quarter of 2015, on the other hand, the advance estimate of 0.2 percent was revised downward to a decrease of 0.7 percent; the most recent estimate for the period was a robust increase of 2.0 percent. (Putting this in context today, 0.1 percent of U.S. GDP – that is, one-thousandth of the GDP – equals about $18 billion.)

These data clearly demonstrate the trade-off between timeliness and more complete and accurate information. It is always important to be mindful that data can be subject to considerable revision. At the same time, policymakers can't afford to ignore the latest numbers in trying to gain an understanding of the state of the economy – for example, to assess whether we are at a turning point in the business cycle and should adjust macroeconomic policy accordingly.

In this case, it is often useful to combine data from multiple sources to gain a more accurate picture of current conditions. The remaining five tips offer examples of how to do this, with the caveat that even these methods can only minimize, not eliminate, the inherent difficulties in measuring the economy.



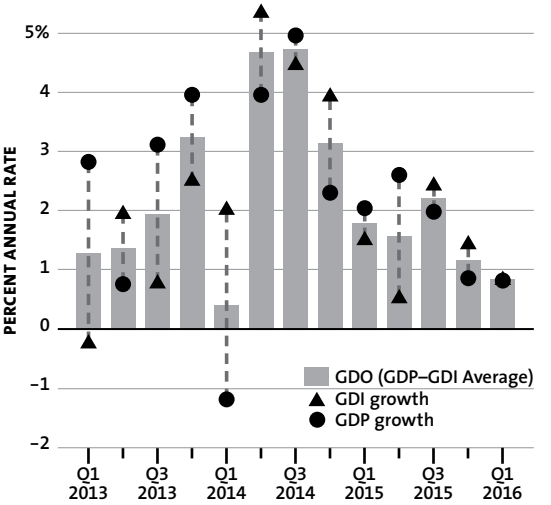
REPORTED GDP GROWTH RATES vary substantially from quarter to quarter. Some volatility is due to true economic fluctuations and some to measurement problems, but it is difficult to figure out whether a startling data point is due to reality or our measurement of reality. For example, it is possible that the economy contracted dramatically in the first quarter of 2014 and then grew rapidly in the next (as Figure 4 shows). But patterns like this seem much more likely to reflect noise in the data.

With economic output, the Bureau of

Economic Analysis publishes an alternative to GDP called gross domestic income (GDI), which aggregates income flows including wages, salaries and business profits. Conceptually, these two measures of output – GDP and GDI – should be equal in a given quarter because the sum of expenditures in the economy should equal the sum of income in the economy. In reality, GDP and GDI nearly always differ, because each measure relies on different data sources and methods of statistical estimation.

It turns out that an equal-weighted average of the two indicators is close to the optimal way to combine them, since the average of GDP and GDI more closely tracks the most up-to-date estimates of GDP growth and is a better predictor of future economic growth than either GDP or GDI alone. The Bureau of Economic Analysis now publishes this average, a concept that the Council of Economic Advisers refers to as "gross domestic output" or GDO, as part of its quarterly data on output. As shown in Figure 5, GDO provides a much more stable reading of the economy than either GDP or GDI. GDO growth shows a similar pattern as GDP growth in the first half of 2014, albeit a less dramatic one than the noisier GDP data alone. And over the last four quarters GDO growth has been relatively steady, avoiding the zigzags that measured GDP growth has undergone.

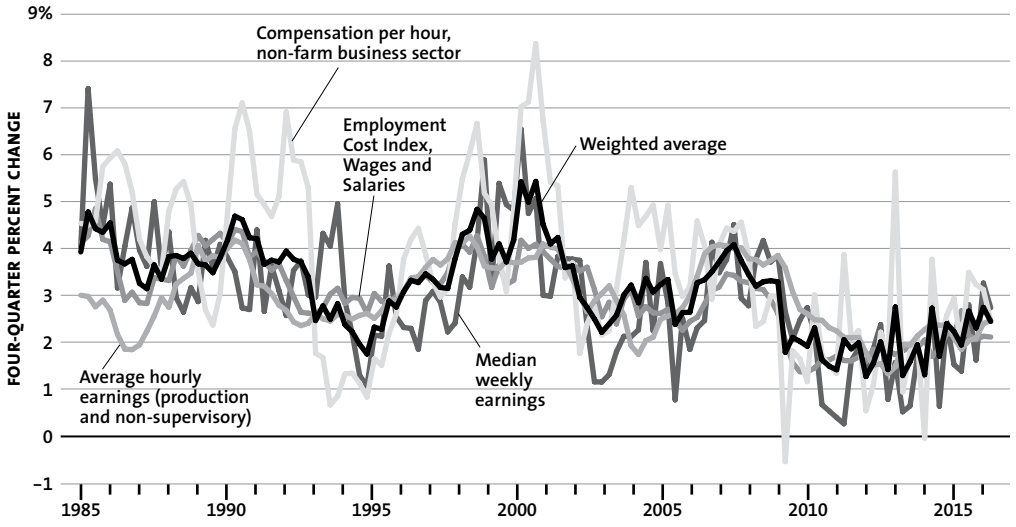
FIGURE 5: GDP AND GDI GROWTH, 2013-2016



SOURCE: Bureau of Labor Statistics, National Income and Product Accounts

COMBINING MULTIPLE INDICATORS can be valuable even when they measure somewhat different concepts. One example of how combining different measures of a similar concept can be useful concerns wages. Statistical agencies publish a dozen-plus measures of wages and labor compensation for the U.S. economy. There are conceptual differences among them, so even if they were measured perfectly they would still track differently. But much of the difference among them is almost surely the result of challenges in statistical measurement (such as sampling error) that are uncorrelated across the different measures. As a result, by combining these

FIGURE 6: MEASURES OF WAGE GROWTH, 1985–2016



NOTE: Weights for average derived from principal component analysis.
SOURCE: Bureau of Labor Statistics; CEA calculations

measures, we can lessen the influence of measurement error and can build a better picture of the underlying trend.

Figure 6 shows four such measures – compensation per hour, average hourly earnings, the employment cost index for wages and salaries and median usual weekly earnings – as well as a weighted average of the four. (The weights are generated through principal component analysis, a statistical technique that extracts common information that may be contained in each series.) Despite the volatility in each measure, the weighted average isolates the consistent story they tell: wage growth in the United States remains below its historical average, but has picked up substantially over the past year. At times, however, growth in this weighted average has differed by 1.5 percentage points or more from the common headline estimate of growth – average hourly earnings for private production and non-supervisory workers – which receives roughly one-quarter weight in the overall index.

EVEN WHEN COMBINING various measures of the same concept to improve an estimate, different aspects of the economy sometimes provide contradictory signals. In this case, it is important not only to estimate the truth in any one measure but to understand the full context of the data available.

In the first quarter of 2014, GDO increased 0.4 percent, while non-farm employment rose by 1.5 percent. (Both figures are annualized rates.) In theory, both of these could be correct – businesses may have stepped up their hiring while workers became less productive, thus decreasing total output. As a matter of accounting, these two concepts are reconciled in the productivity statistics, which show that productivity fell by 3.7 percent at an annual rate in the first quarter of 2014. Measured labor productivity growth is, in fact, extremely volatile, as shown in Figure 1 earlier. That reflects a combination of measurement error in both the numerator (output) and denominator (hours worked) and undoubtedly overstates the true volatility of productivity.

This suggests that, when output and employment are sending diverging signals, the truth is likely somewhere in between – again, implying that combining different measures may be superior to viewing each in isolation. In this case, it is reasonable to put substantially more weight on early estimates of employment growth than on early estimates of output growth, in part because GDP growth is typically subject to larger revisions. Even after a more accurate measure of output like GDO arrives, one should still place more weight on employment growth than output growth.

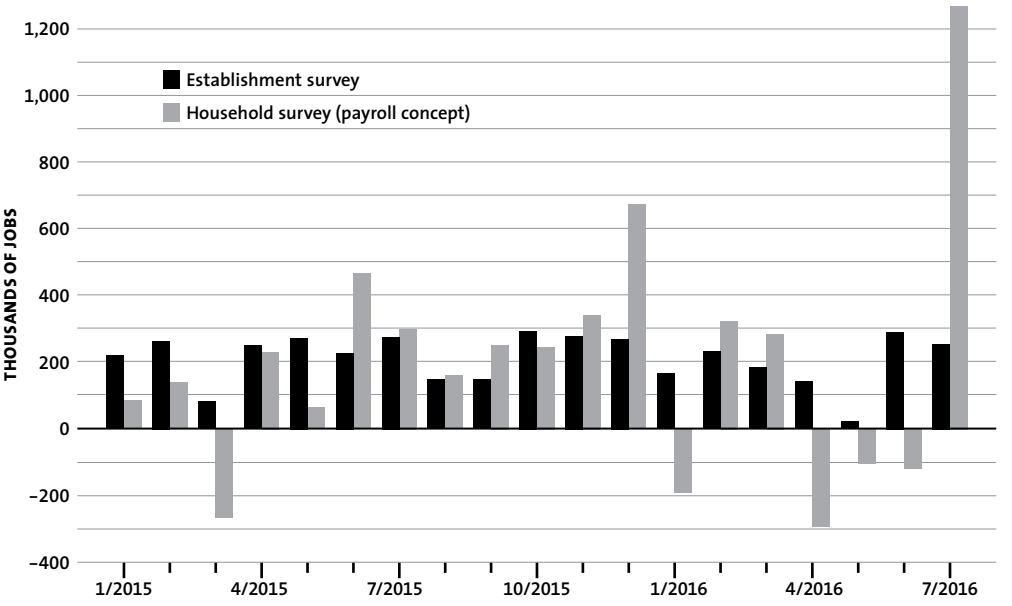
THE BUREAU OF LABOR STATISTICS publishes two measures of job growth every month: the "establishment" or "payroll" survey, which asks employers how many people they have on their payrolls, and the household survey, which asks if individuals are employed or unemployed. Like the different measures of wages, these represent different ways of measuring very similar concepts. As a result, differences between them are more likely to reflect noise than reality – so, in theory, combining them could provide a superior measure of job growth than relying on either one alone.

In understanding whether or how to combine them it is important to note that the household survey in particular is extremely volatile from month to month, as shown in Figure 7, with many instances of sudden spikes in gains (1.27 million jobs added in July 2016) and losses (293,000 jobs lost in April 2016). The establishment survey, on the other hand, can also be somewhat volatile, but does not show nearly the same dramatic month-to-month swings.

This difference in volatility is due to the fact that the establishment survey includes about 440,000 worksites covering about one-third of U.S. employees. In contrast, the household survey is based on only 60,000 households. But the establishment survey is imperfect and suffers from both statistical noise and systematic errors, especially in recording employment gains at new firms that have just come into existence and employment losses at old firms that have closed.

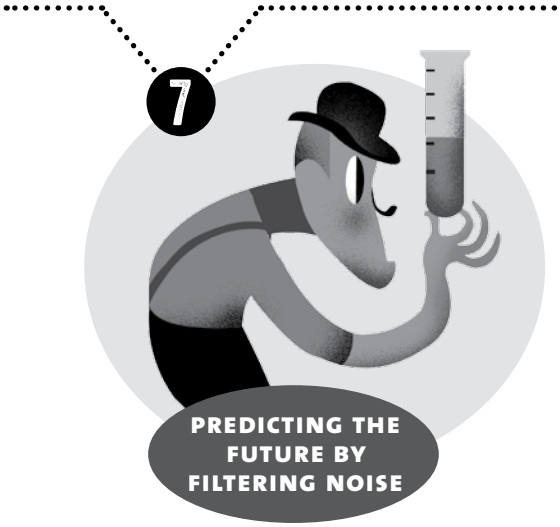
But combining the two does not always lead to more accurate estimates. Although in theory both should contain some information, in practice the household survey is so volatile that it contains practically none. The

FIGURE 7: MONTHLY EMPLOYMENT GROWTH, JANUARY 2015–JULY 2016



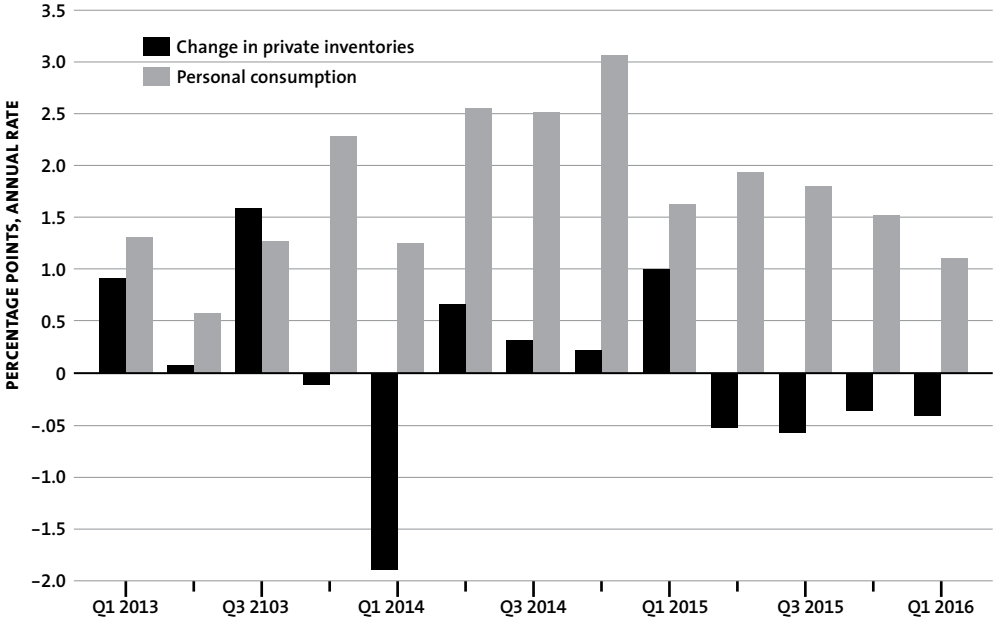
NOTE: Data as of August 9, 2016. **SOURCE:** Bureau of Labor Statistics

optimal weights to combine information from the payroll and household survey is something like 95 to 100 percent of the weight on the former and 0 to 5 percent of the weight on the latter. So, for all practical purposes, the household survey contains virtually no new incremental information about job growth – and it might as well be ignored. Job growth should thus be estimated by the number in the payroll survey.



AS ASSESSED BY GDO GROWTH, the economy was relatively weak in the first quarter of 2014. A more important question at the time was how much of the weakness was transitory (for example, the result of bad weather) and how much was likely to carry forward. Answering this question involves forecasting future economic performance, something that should only be done with great trepidation and humility, given the large uncertainties inherent in the economy and the limitations of our understanding of how it works.

FIGURE 8: CONTRIBUTIONS TO GDP GROWTH



NOTE: Data as of August 9, 2016. **SOURCE:** Bureau of Economic Analysis, National Income and Product Accounts

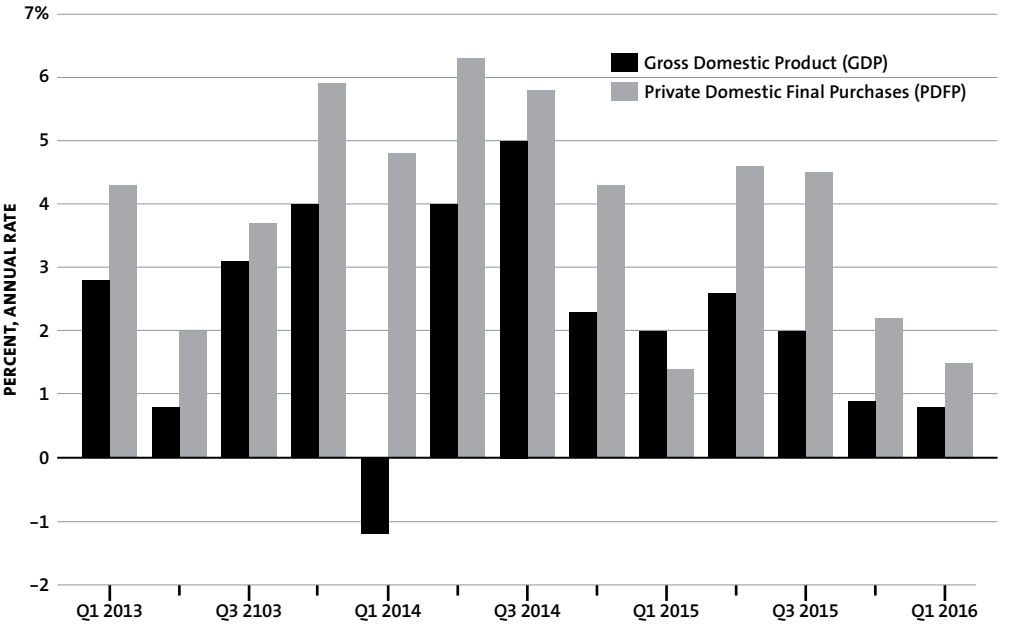
That said, one option is to build an intricate model of the economy using statistical techniques. Another approach is to use judgment – for example, looking at the weather and attempting to assess the impact it could have had on different components of GDP. A third approach is simply to extrapolate from recent performance, a method that, while simplistic, can still generate reasonable forecasts. The question, however, is from which measurements of recent performance we should extrapolate.

As discussed above, economic data are noisy in part because of statistical quirks, but the economy itself is also subject to transitory fluctuations. When we try to measure economic output writ large, we must be mindful that broad aggregates like GDP encompass many different components, some of which contain useful information about the future and some of which do not.

A historical review of the different components of GDP, for example, can give us a better sense of which components are transitory and which tend to be persistent indicators of economic growth. Within GDP, inventory investment bounces around without a clear longer-term trend, as shown in Figure 8, with performance in any given quarter not telling us much about the likely performance in the next quarter. Figure 8 also shows personal consumption, which tells a different story. Personal consumption is about half as volatile as inventory investment and tends to be much more persistent.

One way to make use of this observation is to focus not on the growth rate of GDP as a whole but on the growth rate of personal consumption and fixed investment – a combination called private domestic final purchases (PDFP). Analysis by the Council of Economic Advisers has found those two components of

FIGURE 9: QUARTERLY GDP AND PDFP GROWTH, 2013–2016



SOURCE: Bureau of Economic Analysis, National Income and Product Accounts

GDP to be the most stable, which means that they are the best predictors of GDP over the next quarter or the next year. (GDO is a more accurate measure of what economic performance actually was in the current quarter, complete with true transitory shocks like weather.)

As shown in Figure 9, GDP growth in the first quarter of 2014 was negative, in part because of bad weather. PDFP growth also slowed in the same quarter because of the rough winter, but to a lesser extent than GDP growth did. In the second quarter of 2014, there was a large rebound in GDP growth, but a smaller bounce-back in PDFP because it is a more stable measure of the underlying trend in the economy. In general, PDFP is a more stable, less volatile measure because it contains more of the signal of underlying trends in the economy, while GDP picks up more of the noise. As a result, PDFP is a better predictor of future economic performance than current-quarter GDP.

LAST THOUGHTS

It's pretty clear that the most reliable long-term economic analysis comes from looking across many samples, time periods, measurements and concepts and avoiding putting too much emphasis on any single piece of information. This is difficult when even a casual observer faces a deluge of economic data.

It's important not to overstate the precision we can derive even from a careful reading of these inherently imperfect measures, given the level of uncertainty and volatility in the economy. But by taking a holistic view of economic data and considering each new report in the context of other data as they become available, it is possible to get a much deeper, and more accurate, understanding of what's really happening.