

Rethinking GDP

It may be time to devise a new measure of economic welfare with fewer flaws





Diane Coyle

WHY does economic growth matter? The answer for economists is that it measures an important component of social progress—namely, economic welfare, or how much benefit members of society get from the way resources are used and allocated. A look at GDP per capita over the long haul tells the story of innovation and escape from the Malthusian trap of improvement in living standards that is inevitably limited by population growth.

GDP growth is instrumentally important as well. It is closely correlated with the availability of jobs and income, which are in themselves vital to people's standard of living and underpin their ability to achieve the kind of life they value (Sen 1999).

However, GDP is not a natural object, although it is now everyday shorthand for economic performance. It cannot be measured in any precise way, unlike phenomena in the physical world. Economists and statisticians understand, when they stop to think about it, that it is an imperfect measure of economic welfare, with well-known drawbacks. Indeed, early pioneers of national accounting, such as Simon Kuznets and Colin Clark, would have preferred to measure economic welfare. But GDP prevailed because the demands of wartime called for a measure of total activity. So from the very start, the concept of GDP has long had its critics. But coming up with a better gauge of welfare is easier said than done.

Short-term measure

GDP measures the monetary value of final goods and services—that is, those that are bought by the final user—produced and consumed in a country in a given period of time. The limit of GDP as a measure of economic welfare is that it records, largely, monetary transactions at their market prices. This measure does not include, for example, environmental externalities such as pollution or damage to species, since nobody pays a price for them. Nor does it incorporate changes in the value of assets, such as the depletion of resources or loss of biodiversity: GDP does not net these off the flow of transactions during the period it covers.

The environmental price of economic growth is becoming clearer—and higher. The smog over Beijing or New Delhi, the impact of pollution on public health and productivity in any major city, and the costs of more frequent flooding for which countries are still ill-prepared are all illustrations of the gap between GDP growth and economic welfare. This is why economists and statisticians have been working to introduce estimates of natural capital and its rate of loss (World Bank 2016). When they do, it will be clear that sustainable GDP growth (that enables future generations to consume at least as much as people today) is lower than GDP growth recorded over many years. Getting these new measurements into the mainstream policy debate and reflected in political choices, however, is another matter.

Indeed, GDP ignores capital assets of all kinds, including infrastructure and human capital; it is an inherently short-

term measure. Economic policies aimed at delivering growth have demonstrated the validity of the famous comment of their intellectual architect, John Maynard Keynes: “In the long run we are all dead.”

Seventy years on, the long run is upon us. A broad measure of the sustainability of economic growth, and thus long-term economic welfare, would take account of economic assets as well as the flows counted in GDP: the need to maintain infrastructure or record its depreciation as bridges crumble and roads develop potholes. A true national balance sheet would account for future financial liabilities, such as state pensions. It would also include increases in human capital as more people attain greater education and skill. Economic welfare must be calculated net of such changes in the value of national assets.

Household work

A long-standing criticism of reliance on GDP as the measure of economic success is that it excludes much unpaid work by households. There must be an accepted definition of what is part of the economy and measurable and what is not. Economists call this a “production boundary.” What is within that boundary and what is not inevitably involve matters of judgment. One early debate was whether government spending should be included—on the grounds that it is collective consumption—or excluded—on the grounds that the government is paying for things like roads and security that are inputs into the economy (just like a business expense) rather than consumption or investment goods.

Another key debate concerned how to define goods and services produced—and often also consumed—by households. Home-produced goods such as food were included, because in many countries these can just as easily be bought and sold in the market, but home-provided services such as cleaning and child care were not. Not surprisingly, feminist scholars have always decried the fact that work done mainly by women is literally not valued. Many economists agreed in principle, but the line was drawn partly for reasons of practicality: surveying household services was a daunting task, and these services were seldom purchased in the marketplace.

This of course has changed dramatically in the economies of the Organisation for Economic Co-operation and Development (OECD) since the 1940s and 50s, when the production boundary decisions were made. As more women work for pay, the market for services such as cleaning and child care has grown, and households can and often do switch between performing and buying these services. There is no logical reason not to treat household work like any other work.

The evolution of the digital economy has reignited this old debate, as it is starting to change the way many people work. National accountants have treated government and business as the productive part of the economy and households as non-productive, but the previously relatively clear border between home and work is eroding. More and more people are self-employed or freelance through digital platforms. Their hours

Measuring up

GDP is the monetary value of the total output of goods and services in an economy during a specific period.

Although the definition seems straightforward, deriving GDP is not. First, collecting the data is immensely complicated. There are millions of producers, products, services, and prices.

Moreover, figuring out how much a change in GDP, which is measured in current dollars (or other national currencies), represents a real change in the amount of goods and services available to consumers and how much is due to changes in prices adds another layer of complexity.

If the price of shoes, say, is 5 percent higher than a year ago and GDP registers a 5 percent increase in the value of shoe output, the nominal increase in the shoe component of GDP is an illusion, due to inflation. The actual output of shoes was constant. To determine how much of any, say, year-to-year change in GDP reflects more final output (volume) and how much reflects higher prices (inflation), economists use a technique called deflation.

GDP is a measure of the final goods and services produced in an economy, those that are consumed by people or businesses. **Intermediate goods and services** are netted out in GDP because they are used to produce another good or service. An automobile is a final good. The steel, plastic, and glass, for example, that are used to make it are intermediate products (or inputs).

Three measures

There are three ways to measure GDP. The **expenditure** approach adds up the market value of all spending on final products by consumers, businesses, and government plus exports minus imports. The **production** approach adds up the value of everything that is produced, gross output, then deducts the value of the intermediate products to get net output. The **income** approach adds up everything earned by people and firms—mainly wages, profits, rents, and interest income.

may be flexible, and work can overlap with other activities. In many cases they are using household assets, from computers and smartphones to their homes and cars, for paid work. Many people contribute free digital work such as writing open-source software that can substitute for marketed equivalents, and it clearly has great economic value despite a price of zero.

These developments underline the need for a much better statistical understanding of household activity, yet few countries collect any suitable information on household assets.

Ever-evolving technology

The blurring of boundaries between home and work is not the only way in which technology is making GDP calculation difficult. Many in the technology sector argue that conventional GDP statistics understate the importance of the digital revolution. The pace of innovation has not slowed in areas such as telecommunications, biotechnology, materials, and green energy, they rightly point out—making the lackluster growth and productivity performance of so many advanced economies even more of a puzzle.

For instance, compression technology allows wireless networks to carry more data faster than ever at high quality, and

All three measures theoretically come up with the same value for GDP. But because of difficulties in collecting the source data, the three approaches never give the same value. In many countries, the official GDP is based on the production approach because source data from producers are more comprehensive and accurate.

Price effects

Because the prices of goods and services are collected in current dollars, the so-called nominal GDP is affected by changes in prices and does not necessarily reflect whether or by how much the volume of those goods and services has increased—which is what interests most people and businesses. To see the effects of inflation on the prices of goods and services, economists construct a statistic called an index, which takes account of changes in the price of a good or a service between a base year and the current year. That index is applied to prices to take out the inflation component (or deflate) in current prices.

To return to the shoe example, if the nominal value of shoes rose 10 percent over a year, the nominal GDP for that year would reflect a 10 percent increase in shoe output. If the price of shoes rose 8 percent, then a deflator applied to the shoe price part of GDP would turn that 10 percent nominal increase into a 2 percent real increase (in statistical lingo, the volume of shoes produced rose 2 percent).

Deflators present their own difficulties. The more precise the deflator, the more accurate the real GDP calculation. But there is a sizable drawback. The more precise the deflator, the more information about prices is needed, and collecting price data is costly.

This box draws in part on “Measure Up: A Better Way to Calculate GDP” (IMF Staff Discussion Note 17/02), by Thomas Alexander, Claudia Dziobek, Marco Marini, Eric Metreau, and Michael Stanger.

the price of such innovations as solar energy and genome sequencing has been falling rapidly. Could it be that the statistics are not properly adjusting for quality improvements arising from technology and therefore overstate inflation and understate productivity and growth in real terms?

Official figures in practice incorporate very little quality adjustment to calculate “hedonic” price indices—that is, those that take into account quality improvements. Researchers who have tried to extend hedonic adjustment to a broader range of prices in the information and communication technology sector in the United States have concluded that it makes little difference to the picture of slow productivity growth, in part because there is little US-based information and communications technology manufacturing (Byrne, Fernald, and Reinsdorf 2016).

However, this research has not been extended to the far wider range of goods and services affected by digital transformation, and there are some conceptual questions that need to be resolved. For example, is a streamed music service equivalent to a digital download or buying compact discs, or is it a new good? In other words, is the consumer buying a specific format or simply the ability to listen to music? If the former,

ideally there would need to be a quality-adjusted music price index. In principle, price indices calculate what people have to pay to attain the same level of “utility” or satisfaction from all their purchases, but putting this calculation into practice is not straightforward.

GDP’s shortcomings have become especially obvious in its failure to account for inequality.

Indeed, economists argue that it is impossible to capture all the economic welfare benefits of innovations in GDP, which measures transactions at market prices; there will always be some utility above and beyond that price, labeled “consumer surplus.” Digital goods are no different from previous waves of innovation in this regard. Those who use GDP growth as a measure of economic performance must keep in mind that it has never been a complete measure of economic welfare. For example, the consumer benefits of an important new medicine will eventually always far exceed the market price. This argument, while correct, plays down the possibility of a particularly wide wedge between welfare and GDP today, given digital technology’s effects on business models and consumer behavior.

Inequality matters

GDP’s shortcomings have become especially obvious recently in its failure to account for inequality. The aggregation of individual incomes or expenditures into GDP ignores distributional questions, and equating GDP growth with an improvement in economic welfare assumes that there is no reason for anything other than the status quo distribution. When income distribution did not change much—until the mid-1980s in most OECD countries—ignoring the issue did not matter much. However, thanks in part to Thomas Piketty’s bestselling *Capital in the Twenty-First Century* and in part to the populist movements springing up in many countries, nobody is ignoring distributional questions anymore.

It is possible to adjust GDP to take account of distribution and other nonmarket aspects of economic welfare. Economists have started to debate (once again) specific adjustments. Dale Jorgenson of Harvard University proposes combining distributional information from household surveys with the national accounts (Jorgenson, forthcoming). Charles Jones and Peter Klenow have proposed a single measure incorporating consumption, leisure, mortality, and inequality; their calculations show that this approach closes much of the apparent gap in living standards between the United States and other OECD countries when this is assessed on the basis of GDP per capita (Jones and Klenow 2016).

These measures, extending the standard national accounts approach in a way that at least takes inequality into account, address some of the challenges to gauging GDP, but not all. The debate about how best to measure economic welfare is

intensifying for several reasons. The 2008 global financial crisis and its aftermath are casting a long shadow. Although inequality has begun to diminish in some countries, sluggish growth, debt overhang, and high unemployment in some cases have made for a lackluster recovery and simmering discontent with economic policy that follows business as usual. At the same time, it is hard to ignore the evidence of the environmental cost of past economic growth. The digital revolution and debate about the links between technology and productivity growth—and technology and future jobs—add a subtle twist.

It is easier to express dissatisfaction with current measures than to reach consensus on what should replace GDP. The landmark Stiglitz-Sen-Fitoussi Commission in 2009 recommended the publication of a “dashboard” of economic welfare measures, arguing that its multiple dimensions could not sensibly be reduced to one number. Others argue that a single indicator is essential to have traction in the media and political debate. GDP is set by a slow and rather low-profile international consensus process, so it is hard to imagine any clean break with the current standard unless economic researchers can come up with an approach as compelling in theory and as feasible in practice as GDP, the best-known measure in the framework System of National Accounts.

This might happen. The question is on economists’ research agenda for the first time since the 1940s and 50s. In the United Kingdom, the Office for National Statistics has set up a new research center on economic statistics, launched in February 2017. It is a vitally important debate, given the widespread belief that—as calculated by GDP—recent economic progress has not measured up. Public conversation about economic policy is largely conducted in terms of GDP growth, so the erosion of GDP’s status as a reasonable measure of economic welfare is a serious matter indeed. ■

Diane Coyle is Professor of Economics, University of Manchester, and author of GDP: A Brief but Affectionate History.

References:

- Byrne, David M., John G. Fernald, and Marshall B. Reinsdorf. 2016. “Does the United States Have a Productivity Problem or a Measurement Problem?” BPEA Conference Draft, Brookings Institution, Washington, DC.
- Coyle, Diane. 2015. “Modernising Economic Statistics: Why It Matters.” National Institute Economic Review 234 (November): F4–F8.
- . Forthcoming. “The Political Economy of National Statistics.” In *National Wealth*, edited by K. Hamilton and C. Hepburn. Oxford: Oxford University Press. Working paper version available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2850061
- Jones, Charles I., and Peter J. Klenow. 2016. “Beyond GDP? Welfare across Countries and Time.” *American Economic Review* 106 (9): 2426–57.
- Jorgenson, Dale. Forthcoming. “Within and Beyond the GDP: Progress in Economic Measurement.” *Journal of Economic Literature*.
- Sen, Amartya. 1999. *Development as Freedom*. Oxford: Oxford University Press.
- Stiglitz, Joseph E., Amartya Sen, and Jean-Paul Fitoussi. 2009. *Report by the Commission on the Measurement of Economic Performance and Social Progress*. Paris.
- World Bank. 2016. “Natural Capital Accounting.” Brief, Washington, DC.