

# ONE

## From the Eighteenth Century to the 1930s: War and Depression

**W**arfare is the mother of invention. Many new technologies that end up in use in civilian life have been spurred by the demands of conflict and funded by the military. Among these inventions, ranging from the Internet to Teflon, radar to programmable electronic computers, is Gross Domestic Product. GDP is one of the many inventions of World War II.

Its name sounds as though it should be self-explanatory. *Product*: things that are produced. *Domestic*: at home. *Gross*: nothing deducted, the opposite of *net* (conversely, a cereal packet will give “net weight,” meaning the contents alone, not including the packaging). GDP is just one figure in a full set of accounts for the economy, the national income accounts. We will get to the detail later. To make sense of the idea of GDP, first a brief history of the development of national statistics will help.

## THE EARLY DAYS OF NATIONAL ACCOUNTING

It was an earlier war that prompted the first systematic attempts to measure the whole of the economy. In 1665, a British scientist and official, William Petty, produced estimates of the income and expenditure, population, land, and other assets of England and Wales, with the aim of assessing the country's resources to fight a conflict and finance it through taxes (it was the now little-known Second Anglo-Dutch War, which lasted from 1664 to 1667). Petty wanted to prove not only that the country could bear a higher burden of taxes but also that it was capable of taking on its powerful neighbors, Holland and France.<sup>1</sup> There was no need for it to win more land or increase the size of the population to ensure victory, because the available land and capital and labor could be used to better effect. This was a significant economic insight. Also significant was Petty's introduction of the tool of double-entry bookkeeping to keep records for the nation as a whole. Another early set of estimates, by Charles Davenant in 1695, had the title *An Essay upon the Ways and Means of Supplying the War*, making his aim perfectly clear. The word *statistics* has the same origin as *state*, and originally referred to the collection of figures concerning the state, specifically taxes. It proved to be a major advantage for England to have consolidated national income statistics, enabling calculations about the scope for increased output and tax revenues, when its larger and seemingly more powerful neighbor France lacked such information. Not until 1781 did the French king have similar strategically important economic and financial data, when the finance minister Jacques Necker delivered a famous *compte rendu au roi*, or report to the king, on the strength of the French economy. It enabled the king to raise new loans

but did not, of course, help him avert the French Revolution in 1789.

Throughout the eighteenth century a number of successive statistical pioneers built on these first British attempts, although each was measuring slightly different things. The concept of “national income” may seem clear enough, but measuring it in practice means choosing what to include and exclude, which is surprisingly fuzzy. Unlike in our own time, there was no standardization, no commonly agreed definition; and what was measured was not at all like modern GDP. What these early national accounts had in common was the general idea that the national income depended on how much was available to spend now and how much remained for increasing the national stock of assets.

This framework evolved over the decades.<sup>2</sup> Later authors emphasized different aspects of the economy. Some—among them the novelist and pamphleteer Daniel Defoe—thought that the key to the nation’s prosperity was increasing trade, both overseas and within the country. At another time, the debate in coffeehouses and pamphlets centered firmly on the national debt, the figures for which the government published frequently between the late seventeenth and late eighteenth centuries. Once again, financing warfare was the motivation.

Then came a substantial intellectual innovation. In *The Wealth of Nations* (published 1776), Adam Smith introduced the distinction between “productive” and “unproductive” labor. An anonymous author had written in 1746, “What I mean by National Income is, all the whole body of our People get or receive from Land, Trade, Arts, Manufactures, Labour, or any other way whatsoever; and by Annual Expence I mean, the whole that they spend or consume.” Yet in Adam Smith’s definition thirty years later, the “whole body of our

People” did *not* count. Only those involved in the making of physical commodities, agriculture and industry, would count toward national income. The provision of more *services* was a cost to the national economy, in his view. A servant was a cost to his employer, and did not create anything. Importantly, money spent on warfare or the interest on government debt was also being used unproductively. The nation’s wealth was its stock of physical assets less the national debt. National income was what derived from the national wealth. According to Benjamin Mitra-Kahn, “*The Wealth of Nations* introduced a new idea of the economy, and through the effort of Adam Smith’s students and admirers, it was adopted almost instantly.”

In Smith’s own words:

There is one sort of labour which adds to the value of the subject upon which it is bestowed: There is another which has no such effect. The former, as it produces a value, may be called productive; the latter, unproductive labour. Thus the labour of a manufacturer adds, generally, to the value of the materials which he works upon, that of his own maintenance, and of his master’s profit. The labour of a menial servant, on the contrary, adds to the value of nothing. . . . A man grows rich by employing a multitude of manufacturers: He grows poor, by maintaining a multitude of menial servants.<sup>3</sup>

The idea of a distinction between productive and unproductive activity, adopted by Adam Smith, dominated economic debate and measurement until the late nineteenth century. Karl Marx echoed it, and it remained the basis for measuring the centrally planned economies until the collapse of communism after 1989. For example, the Soviet Union’s economic statistics counted material output and largely ignored service activities; yet by the late 1980s, these accounted for about two-

thirds of GDP in the Western capitalist economies, so it was a large omission.

Still, this way of thinking about the national economy in terms of material production was generally adopted in the nineteenth century, until it too was overturned. Then the new generation of “neoclassical” economists (in contrast to “classical” economists such as Adam Smith) discarded the distinction between productive and unproductive activities. Alfred Marshall, as titanic a figure as Smith in the history of economic thought, said firmly: “Wealth consists of material wealth and personal or non-material wealth.” Services were to be included in the definition of national income. The work done in the late nineteenth and early twentieth century to measure the economy in the wake of Marshall’s decree in his 1890 book *Principles of Economics* has been described as a “first phase” of national income accounting.<sup>4</sup>

### THE BIRTH OF MODERN NATIONAL ACCOUNTS

This quick dip into the early history of national income accounts and the forerunners of GDP shows that the definition of “national income” was not precise or fixed. How it was interpreted depended on the intellectual climate and on the political or military needs of the moment, and so the definition changed over time. Some economists have concluded that the pre-twentieth-century measurement of the economy was not all that serious. Angus Maddison, who spearheaded the extraordinary achievement of constructing GDP statistics for the world dating back to AD1000, wrote, “Economic growth was much slower before the 19th century and therefore seemed irrelevant or uninteresting.”<sup>5</sup> He added, a bit sniffily: “Although there was a proliferation of national income estimates, there

was little improvement in their quality or comparability. They provided little help for serious analysis of economic growth, and there were significant differences in their coverage and methodology.” The early work was certainly not consistent over the years, nor consistent with our modern definitions. But the opposite interpretation seems likely: from the nineteenth century, people were starting to reconsider how to measure the economy precisely because it was starting to grow, thanks to the Industrial Revolution and the dawn of capitalism.

The definitions we use now date back to two seismic events in modern history, the Great Depression of the 1930s and World War II (1939–1945).<sup>6</sup>

Following the publication of Alfred Marshall’s *Principles of Economics*, a number of researchers had already set about new efforts to improve the collection of statistics and the measurement of national income. In the United Kingdom, the most successful was due to Colin Clark, who throughout the 1920s and 1930s calculated national income and expenditure, for the first time on a quarterly rather than an annual basis, and with a new degree of care and thoroughness. For example, he provided detailed splits of production and expenditure into different categories and published thorough accounts of the government’s finances, too. He discussed how to adjust the figures for inflation, and also the distribution of income among different categories of people. Clark was appointed in 1930 to provide statistics to the newly created National Economic Advisory Council, the first body ever created by the British government to provide formal economic advice. The experience of the Depression created this demand for statistics that might help the government figure out how to bring to an end the unprecedented economic slump.

Across the Atlantic, in the United States, Simon Kuznets had a similar motivation. The government of Franklin Del-

and Roosevelt wanted a clearer picture of the state of an economy trapped in a seemingly endless depression. The National Bureau of Economic Research was requested to provide estimates of national income. Kuznets, who later won the Nobel Memorial Prize in Economic Science for this work, took on the task of developing Clark's methods and applying them to the U.S. economy. He was a meticulous collector and assembler of data, paying careful attention to the circumstances in which different statistics were gathered, and what their flaws might therefore be.<sup>7</sup> His first report, submitted to Congress in January 1934, showed that America's national income had been halved between 1929 and 1932. Even in those depressed times the report was a bestseller, at twenty cents a copy, and the first print run of forty-five hundred copies quickly sold out.<sup>8</sup> President Roosevelt cited the figures in announcing the new Recovery Program and used updated figures (running up to 1937) to send a supplemental budget to Congress in 1938. As one survey of the history of national accounting points out, having national income estimates for the whole economy made a huge difference to the scope for policy. President Herbert Hoover had made do with the incomplete picture painted by industrial statistics such as share price indexes and freight car loadings. This information was less compelling, as a call to action, than an authoritative figure showing the halving of the whole of national economic output in the space of just a few years.

Kuznets, however, specifically saw his task as working out how to measure national economic *welfare* rather than just *output*. He wrote:

It would be of great value to have national income estimates that would remove from the total the elements which, from the standpoint of a more enlightened social philosophy than that of an acquisitive society represent dis-service rather than service. Such

estimates would subtract from the present national income totals all expenses on armament, most of the outlays on advertising, a great many of the expenses involved in financial and speculative activities, and what is perhaps most important, the outlays that have been made necessary in order to overcome difficulties that are, properly speaking, costs implicit in our economic civilization. All the gigantic outlays in our urban civilization, subways, expensive housing, etc., which in our usual estimates we include at the value of the net product they yield on the market, do not really represent net services to the individuals comprising the nation but are, from their viewpoint, an evil necessary in order to be able to make a living.<sup>9</sup>

These observations prefigure some of the criticisms made of GDP in our own time: GDP definitely does not attempt to measure welfare or well-being (a subject picked up again in chapters 5 and 6).

With this aim, in fact, Kuznets was out of tune with his times. Welfare was a peacetime luxury. This passage was written in 1937, when his first set of accounts was presented to Congress. Before long, the president would want a way of measuring the economy that did indicate its total capacity to produce but did *not* show additional government expenditure on armaments as reducing the nation's output. The trouble with the prewar definitions of national income was precisely that as constructed they would show the economy shrinking if private output available for consumption declined, even if the government spending required for the war effort was expanding output elsewhere in the economy. The Office of Price Administration and Civilian Supply, established in 1941, found that its recommendation to increase government expenditure in the subsequent year was rejected on this basis. Changing the definition of national income to the con-

cept of GDP, rather than something more like Kuznets's original proposal, overcame this hurdle.

There was a heated debate between Kuznets and other economists, especially Milton Gilbert of the Commerce Department, about the right approach. The discussions were highly technical but the underlying issue was profound: what was the meaning of economic growth and why were statisticians measuring it? Gilbert and his colleagues were clear that the aim was to have a measurement that was useful to the government in running its fiscal policy. As one of the pioneers of GDP put it, rather blandly: "It will be convenient if the incomings and outgoings of public authorities in the provision and organization of common services such as defence, justice, education and public health are thought of as consolidated in the consumption box, being in fact nothing more than agency activities for the body of consumers as a whole."<sup>10</sup> An official U.S. history of the national income accounts describes it this way:

Before GNP [Gross National Product] was made available, projected defense expenditures were sometimes erroneously subtracted from projected national income, producing a residual that was interpreted as the amount of production left for non-war goods and services. . . . The assessment was overly grim because national income fell short of the total market value of goods and services produced, of which defense spending was a component. . . .

By including all government purchases as part of national products, the GNP statistics established the role of national government in the economy as that of an ultimate consumer, that is as a purchase of goods and services for final use.<sup>11</sup>

The first American GNP statistics were published in 1942, distinguishing between the types of expenditure, including

by government, and permitted economists to see the economy's potential for war production. "The inclusion of business taxes and depreciation [in GNP measured at market prices] resulted in a production measure that was more appropriate for analysis of the war program's burden on the economy."<sup>12</sup> Kuznets was highly skeptical: "He argued that Commerce's method tautologically ensured that fiscal spending would increase measured economic growth regardless of whether it actually benefited individuals' economic welfare."<sup>13</sup> In the policy tussle in Washington, Kuznets lost and wartime realpolitik won.

This decision was a turning point in the measurement of national income, and it meant that GNP (or later GDP) would be a concept strikingly different from the way the economy had been thought about from the dawn of modern industrial growth in the early eighteenth century until the early twentieth century. For two centuries, "the economy" was the private sector. Government played a small role in economic life, and featured mainly because it looked to raise taxes to pay for wars. Its role expanded steadily over the centuries, however. In Victorian times this began to extend to the provision of other services, those we take for granted now such as roads and water as well as the historic government roles of defense and justice. By the time the wartime economists developed the modern concept of GDP, government was already a far greater presence than it had been. Subtracting defense spending from the older conception of national income would have wrongly given the impression that the war effort was going to involve a huge sacrifice in private consumer spending. There is, of course, a world of difference between a monarch extracting tax revenues to wage war and a democratic government pooling citizens' incomes to provide services and social security. One aspect of this democratic transition was the switch

to conceiving of government as adding to national income rather than subtracting from it. However, the importance of wartime necessity in shaping the definition should not be underestimated. The pattern of growth before and after 1945 would have looked very different if government spending had been disregarded as before in the definition of total economic activity.

The United Kingdom, already since 1939 at war with Germany and its allies, had earlier reached the same conclusion as the U.S. officials. Colin Clark's approach was overtaken and extended when the brilliant and influential John Maynard Keynes published in 1940 the pamphlet *How to Pay for the War*. He fulminated in this about the inadequacy of the statistics available to him for calculating what the U.K. economy could produce with the available resources, what would be required for mobilization and conflict, what would be left over for people to consume—and how much their living standards might need to fall. Planning for the war effort in particular needed much better statistics on how much was produced by individual industries, using what materials. Keynes wrote, "Every government since the last war has been unscientific and obscurantist, and has regarded the collection of essential facts as a waste of money."<sup>14</sup>

Other countries were separately developing the concept and measurement of GDP during the 1930s. Holland was another pioneer, as were Germany and the Soviet Union. But the motivating force of being at war should not be underestimated. Wesley C. Mitchell, the director of the National Bureau of Economic Research, said: "Only those who had a personal share in the economic mobilization for war could realize in how many ways and how much estimates of national income covering 20 years and classified in several ways facilitated the World War II effort."<sup>15</sup>

A senior British Treasury official, Austin Robinson, was so impressed by Keynes's argument in *How to Pay for the War* that he commissioned two young economists, Richard Stone and James Meade, to develop what became the first modern set of national accounts and GDP. These were published with the U.K. government's budget of 1941. Keynes did not have an official post, but was given an office in the Treasury and oversaw the work and the subsequent establishment of the Central Statistical Office, a new official statistical agency. In 1984, Stone was awarded the Nobel Memorial Prize in Economic Science for his contribution to developing GDP and the national accounts (Meade had won the award earlier for his work on trade theory). Stone went on to be particularly influential in the postwar coordination and standardization of the definitions and measurements of GDP. This began as a discussion between the British and American experts. In May 1946, a Committee of Statistical Experts met at Hunter College in New York to draw up recommendations for collecting national statistics on behalf of the United Nations.

The planning required during the conflict continued to be needed in the period of rebuilding after the end of World War II. George Marshall, the then U.S. secretary of state, used a speech at Harvard University on 5 June 1947 to announce U.S. support for postwar reconstruction. He said:

Aside from the demoralizing effect on the world at large and the possibilities of disturbances arising as a result of the desperation of the people concerned, the consequences to the economy of the United States should be apparent to all. It is logical that the United States should do whatever it is able to do to assist in the return of normal economic health in the world, without which there can be no political stability and no assured peace. Our policy is directed not against any country or doctrine but against hunger, poverty, desperation and chaos. Its purpose should be

the revival of a working economy in the world so as to permit the emergence of political and social conditions in which free institutions can exist.<sup>16</sup>

The administration of President Harry Truman lived up to this vision and provided total aid worth an estimated \$148 billion (in 2004 dollars) from 1946 to 1952.<sup>17</sup> The devastated countries of Europe depended heavily on Marshall Aid to enable them to survive and rebuild. Throughout that period everything was in short supply and tracking the use of resources was essential. Before long, the United Nations took on the responsibility for setting international standards of measurement in what is now known as the System of National Accounts (SNA).

Once available, these statistics on the whole economy found another use. Keynes had wanted to have the figures available for the purpose of wartime planning. But he had also published just before the war his massively influential book *The General Theory of Employment, Interest and Money*. At the heart of this classic economic text lies a theory about the relationship between different economic variables, including, in addition to national income, personal consumption, investment and employment, interest rates, and the level of government spending. The theory set out links between the tools the government had available and the size of the economy. It became the basis for a more interventionist approach to government economic policy from the 1940s onward, using both fiscal policy (the level of tax and spending) and monetary policy (the level of interest rates and availability of credit) to target a higher and less volatile rate of growth for the economy. The use of these tools was developed more fully by other economists after Keynes's early death in April 1946. Postwar policymakers still bore the scars of the Great Depression and pounced on the economic theories of Keynes

and his successors as a means of averting a repetition of that crisis. Crucially, the development of GDP, and specifically its inclusion of government expenditure, winning out over Kuznets's welfare-based approach made Keynesian macroeconomic theory the fundamental basis of how governments ran their economies in the postwar era. The conceptual measurement change enabled a significant change in the part governments were to play in the economy. GDP statistics and Keynesian macroeconomic policy were mutually reinforcing. The story of GDP since 1940 is also the story of macroeconomics. The availability of national accounts statistics made demand management seem not only feasible but also scientific.

This sense of control was enhanced by the parallel development of methods for using national accounts statistics to estimate econometric “models” of the economy. The pioneer here was a Dutch economist (and the first winner of the Nobel Memorial Prize in Economic Science), Jan Tinbergen, his country having been almost as rapid in adopting GDP as the United Kingdom and the United States. A macroeconomic model is a series of equations representing certain relationships, for example between interest rates and investment, or consumer spending and incomes. Econometric techniques are the statistical methods used to estimate these relationships on the basis of past statistics—for example, that consumers will spend 40 percent of an increase in personal income. The models estimated on *past* averages can then be used to predict what might happen in the future, especially if the government changes some of its policies, say, increasing personal taxes. For example, a one-million-dollar increase in government spending (or cut in taxes) will put more disposable income in the pockets of taxpayers, who will spend some of the extra on goods and services. The people benefiting from that additional business will have higher incomes them-

selves, and spend more in turn. A key question is how big the final increase in GDP will turn out to be. That depends on other factors: how much extra actually gets spent rather than saved; how much interest rates rise because there's more demand for borrowing; how much is spent on imports rather than goods and services produced at home; how much inflation rises because the additional demand exceeds what is supplied in the short term. As economists would put it, the additional government spending can either "crowd in" or "crowd out" private spending. In the former case, the fiscal "multiplier" exceeds one, and in the latter case is less than one—some estimates have it actually being negative. The multiplier is a measure of how much GDP changes compared to the change in government spending (or tax revenues).

Although Keynes himself was highly skeptical about econometric models, they became a key tool in the more interventionist approach to government economic policy that prevailed from the late 1940s until the economic crisis of the late 1970s. In fact, the number of models proliferated, and a new forecasting industry was created by pioneers such as Otto Eckstein, the founder of Data Resources, Incorporated (DRI).<sup>18</sup> We are now awash with macroeconomic models and forecasts, published by official agencies and central banks, by investment banks, by think tanks and researchers, as well as by commercial forecasters such as DRI's successors. Indeed, the idea of the economy as a machine, regulated by appropriate policy levers, took firm hold. Such firm hold that the engineer-turned-economist Bill Phillips built an actual machine showing the flow of income in the economy and the routes through which government policy could increase that flow (see figure 1). A few of the machines remain in universities, as curiosities—but the "engineering" mindset still has a firm grip on economic policy.



Figure 1. Phillips Machine, courtesy of the Library of the London School of Economics and Political Science, IMAGELIBRARY/6

For macroeconomic models are still in widespread use, even though the earlier illusion of precise control ought to have been shattered by events over the course of the decades since the 1940s. Governments after all still need to try to forecast what the effect of their interventions and policy changes might be. They are now far more complicated and subtle than earlier generations of macroeconomic models (partly because the economy has become more complicated), and importantly include the influence of expectations about the future on current links between economic variables. Nonetheless, the financial and economic crisis since 2008—not predicted by mainstream economic forecasters—has led to an active debate about whether the approach of aggregating individual behavior and assuming stable links between the aggregate measures (that is, the statistics defined in the national accounts) is a valid exercise. Debate rages in particular about the multiplier, because the issue of whether extra government spending or tax cuts (a “fiscal stimulus”) will boost GDP growth turns on its size. If it is greater than one, a stimulus of extra government spending will help growth, while austerity measures will hurt it. Its actual size is hotly contested among macroeconomists, especially in the context of the present political debate about how much “fiscal stimulus” the government should be applying to get the economy growing faster. There is an unsurprising alignment in the “multiplier wars” between macroeconomists’ answer to the technical question about the size of the multiplier and their political sympathies. In early 2013, after some years of austerity measures cutting government spending in European countries and Japan, the IMF’s chief economist concluded that, contrary to the Fund’s earlier official view, the multipliers on fiscal policy in the early years of the crisis were substantially above one, or in other words that austerity measures had done more harm

than good to short-term GDP growth.<sup>19</sup> The paper also makes it clear, however, that multipliers vary among countries and over time (although estimates are typically greater than one), so the mechanical approach to macroeconomic modeling that developed alongside the creation of today's GDP and national accounts figures remains in question.

## THE NATURE OF GDP

It will be clear by now that the ambition of measuring national income has a long history, with correspondingly many changes in how people have thought about it. As Richard Stone put it, national income is not a “primary fact” but an “empirical construct”: “To ascertain income it is necessary to set up a theory from which income is derived as a concept by postulation and then associate this concept with a certain set of primary facts.”<sup>20</sup>

There is no such entity as GDP out there in the real world waiting to be measured by economists. It is an abstract idea, and one that after a half century of international discussion and standard-setting has become extremely complicated. The manuals for statisticians run to hundreds of pages, and it takes a substantial investment of time and effort to understand the national accounts in any detail. Now, though, it's time to set out the basics.

### *What Is It? Definitions*

It is surprisingly hard to write down definitions of GDP that do not assume some prior knowledge. So this section will seem complicated for GDP novices, and yet will be considered laughably oversimplified by national accounts experts.

Understanding GDP is a bit like a video game with increasing levels of difficulty.

The system for measuring GDP and its components has steadily become more and more complicated, too. This is because of the increasing sophistication of the statistical methods used, and because of the increasing complexity of the economy itself. For example, a growing proportion of the economy consists of services, whose output is inherently harder to measure than the output of, say, tractors or cotton fabric. The first UN guide to the SNA, which all countries are supposed to follow, was published in 1953 and had fewer than fifty pages. The 2008 SNA document has 722 pages. A widely used commentary on the SNA has four hundred pages.<sup>21</sup> The community of national statisticians with a command of all this detail is small. In other words, very few people indeed truly understand how the regularly published GDP figures are constructed—this excludes many of the economists who comment on GDP. So, take a deep breath before embarking on the next few pages.

To start with the basics, GDP can be measured in three ways, in principle equivalent to one another. You can add up all the *output* of the economy, all the *expenditure* in the economy, or all the *incomes*. Table 1 shows these three and their components, including their shares in the U.S. economy in 2005.

The rest of this section concerns GDP, but before continuing, it should be noted that another way of defining an economy's total output is Gross *National* Product. GDP counts all the economic output generated within the nation's boundary. GNP counts all the economic output generated by national entities, some of it occurring overseas. In other words, the main difference between the two is that GNP also includes output or income from overseas. For a few small countries (such as Ireland and Luxembourg), the difference is large.

Table 1: Three Ways to Measure GDP

|                                                                      | 2005 share<br>(percent) |
|----------------------------------------------------------------------|-------------------------|
| <b>I. Value-added (or production) approach</b>                       |                         |
| Gross output (gross sales less change in inventories)                | 183.5                   |
| <i>Less:</i> Intermediate inputs                                     | 83.5                    |
| <b><i>Equals:</i> Value added for each industry</b>                  | <b>100.0</b>            |
| <b>II. Income (by type) approach</b>                                 |                         |
| <i>Sum of:</i> Compensation                                          | 56.6                    |
| Rental income                                                        | 0.3                     |
| Profits and proprietors' income                                      | 17.6                    |
| Taxes on production and imports                                      | 7.4                     |
| <i>Less:</i> Subsidies                                               | 0.5                     |
| Interest, miscellaneous payments                                     | 5.5                     |
| Depreciation                                                         | 12.9                    |
| <b><i>Equals:</i> Total domestic incomes earned</b>                  | <b>100.0</b>            |
| <b>III. Final demand (or expenditures) approach</b>                  |                         |
| <i>Sum of:</i> Consumption of final goods and services by households | 70.0                    |
| Investment in plant, equipment, and software                         | 16.7                    |
| Government expenditures on goods and services                        | 19.0                    |
| Net exports of goods and services (exports – imports)                | –5.7                    |
| <b><i>Equals:</i> Final sales of domestic product to purchasers</b>  | <b>100.0</b>            |

Source: J. Steven Landefeld, Eugene P. Seskin, and Barbara M. Fraumeni, "Taking the Pulse of the Economy: Measuring GDP," *Journal of Economic Perspectives* 22, no. 2 (2008): 193–216.

For most it is not. GDP has largely displaced GNP since the start of the modern era of globalization, although arguably economists should pay more attention to divergences between the two. GDP is also a "gross" measure in that it does not make any adjustment for the depreciation of assets (in other words, the wear and tear that reduces their value over time)—deducting this would give net domestic product. In

### The Circular Flow

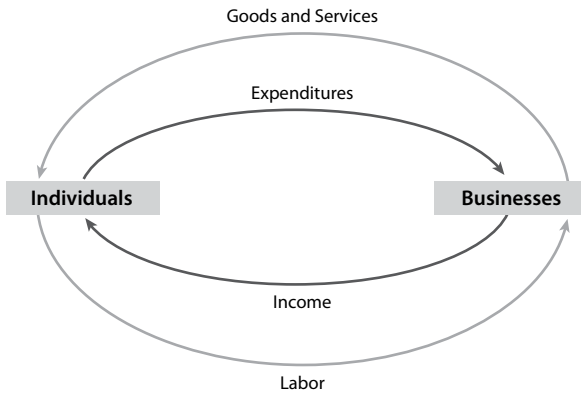


Figure 2. The Circular Flow

some ways this is a more interesting measure (the reasons are covered in chapter 6), but it is rarely referred to in everyday discussions about the economy.

Back to GDP.

Sometimes it is described as a mystical-sounding “circular flow,” as illustrated in figure 2 (and built into the pipes and tubes of the Phillips Machine).

What this means is that the national accounts, like any accounts, add up on each side of the books. One consumer’s spending is a business’s sales revenue—when added up for the whole economy, these corresponding flows of money have to balance.

The approach to measuring GDP most familiar from the newspapers and TV is the expenditures approach. Commentators talk about consumers opening their wallets or businesses reining back investment. The equation

$$\text{GDP} = C + I + G + (X - M)$$

—consumer spending plus investment spending plus government spending plus exports less imports (the trade surplus or deficit)—will be known to anyone who has taken an introductory economics course.

So the idea seems simple. GDP is the sum of all that is spent in the national economy. Expenditures are divided into several categories. Following Keynes, these are C, consumption (by private individuals or households); I, investment (by companies); G, government expenditure (on goods and services—but not transfer payments such as welfare or pensions); and  $X - M$ , exports less imports.<sup>22</sup> Each can be further divided into subcategories, such as spending on food, or investment in buildings, or government expenditure on education.

The practical reality is very complicated, a matter of finicky attention to detail. For example, what does measuring C mean in practice? To quote from one explanatory article, “The commodity flow method starts with total sales (or shipments) by producers of final goods and services. Then using this estimate of final sales, the bureau adds (a) transportation costs (b) wholesale and retail trade margins (c) sales taxes and (d) imports. It then deducts (e) changes in inventories (f) exports (g) sales to business (because these are intermediate goods) and (h) sales to government. The method produces consistent estimates of the value of final sales to consumers.”<sup>23</sup> And so on.

Then there are fuzzy boundaries—for example, should a consumer’s purchase of a car she will use for ten years be counted as “consumption” when a company’s purchase of software it will use for only two years is classed as “investment”? The change in the amount of inventory or stock held by businesses will have to be included, even though it is likely to

change in response to events rather than deliberate business choice. Some elements of expenditure are estimates of the value of services not purchased directly: the value of living in an owner-occupied house is one example.

The approach of adding up everything produced in the economy (the value-added approach) is even simpler to set out, as in table 1, but again in practice requires a lot of careful calculation of the separate elements. The total adds up everything that is produced in the economy, all the goods and services. Almost every business will use in its production the output of other businesses, however. So to avoid double counting, these purchases of “intermediate” goods have to be excluded, netted off from final sales.<sup>24</sup> At first this was a tricky matter of patching together different sources of data. But in the 1950s, Wassily Leontief (another Nobel Memorial Prize-winning economist) came up with the idea of input-output tables that tracked the sale and purchases of intermediate goods through the economy to calculate the “value added” in production. From the mid-1960s this method has been used to publish the production side of the national accounts. (In the USSR, Gosplan, the central planning agency, adopted the same approach, publishing the first Soviet input-output tables in 1959.)

Another important practicality is to adjust for normal seasonal fluctuations. It is not helpful to know that consumers spent more in the October–December quarter than in the July–September quarter, as the holidays alone will mean higher spending every year. We want to know whether this year’s increase was unusually large or small. So there are statistical techniques for “seasonal adjustment” to give figures that smooth out the normal ups and downs over the year. The figures analyzed by economists and pronounced on by the media are all these seasonally adjusted ones. The statistical

bureau takes the initial figures and adjusts them for “normal” seasonal patterns. Abnormal changes in the pattern, such as an unusually hot summer, or a holiday falling on a Thursday so everyone takes a long weekend, can play havoc with the adjustment process.

All of the transactions included in GDP are measured at the price paid for them in the market—when these are available. Government spending, for example, is by definition not in a marketplace, so it has to be valued either by how much the government pays in salaries to the people working in the service, or how much a similar service would cost if purchased privately. An alternative is to measure GDP at factor cost, which adjusts the market price paid by deducting VAT or sales tax and adding in any government subsidy that reduced the price. The “factor cost adjustment” is the gap between the two GDP totals.

As noted a few pages back, the “gross” part of GDP indicates that expenditure is not adjusted for spending on wear-and-tear or obsolescence. This affects some household purchases but more so business investment in assets that need maintenance and repair or need replacing after time. An estimate of depreciation can be deducted from GDP to give net domestic product.

The “accounting” in the national accounts means that the sum of all spending in the economy must, by definition, add up to the sum of all incomes in the economy. These are classed by category—income from employment, from self-employment, from dividends and interest, business profits, income from overseas, and so on. In reality, as expenditures and incomes are collected from completely different sources of data, and a wide range of them too, they will never quite match. The statistical mismatch can be large from time to time. The U.S. and U.K. official statistics publish the discrepancy.

So far, this has been about “nominal” measures, that is, dollar or pound sterling amounts. For the purposes of economic policy, the split between inflation and “real” growth is needed. Achieving higher (nominal) GDP growth through inflation alone would be a bad sign about economic management; that was what happened in the mid-1970s when in many countries governments responded badly to big hikes in the price of oil and got the mix of slow or negative real GDP growth and high inflation labeled “stagflation.” Nominal GDP continued to grow even though living standards were falling and unemployment was rising. So, to calculate real GDP statisticians have to collect data on prices and combine that into a general price index, the GDP deflator.

There are many different ways of creating a price index and calculating inflation. It will be no surprise to learn that inflation-adjustment is very complicated, perhaps one of the most challenging of all the methodological statistical issues. The calculation involves prices and quantities of each component element in the base year and the later year, often weighted depending on the different proportions of each item sold in either the base year or later year. The prices of all the goods and services in the economy are combined according to how much of each is sold. The “weight” given to each item therefore reflects how significant it is in that year’s economic activity. The resulting number is “rebased” to be equal to 100 in a chosen year. The next year’s price index is calculated applying the same weights to the increase in each price between one year and the next. The result will be a number like 102.5 or 104.3 (if it is less than 100, the price level has declined). This is used to divide into the later year’s money GDP level, giving a “real” GDP level in terms of the base year’s dollars or pounds. The importance of this adjustment will crop up again later.

I hope this sounds moderately straightforward. Unfortunately, there are dozens of different ways of calculating the formula, all with different results.<sup>25</sup> So, although we definitely want to make this adjustment for inflation to measure real economic growth, the choice of technique can lead to strikingly different “real” conclusions.

### *Is Africa Poor?*

You might think that this is a straightforward question to answer; after all, the concepts “Africa” and “poverty” are closely linked in the minds of Westerners. It is harder than it seems, however, and this shows why the horrible technicalities involved in constructing GDP statistics, described briefly in the previous section, matter a lot. Specifically, take the question of whether or not Ghana is a poor country. Aid organizations use a threshold in terms of real GDP per capita set by the World Bank to designate whether a country is “low-income” or “middle-income,” and this in turn determines the kind of assistance it gets in aid and cheap loans. Until November 2010, Ghana was considered “low-income,” that is, a poor country. But between 5 and 6 November 2010, its GDP increased by *60 percent* overnight, turning it officially into a “low-middle-income” country. The reality had not changed, but the GDP statistics had, because the country’s statistical agency had updated the weights used in calculating the price index, and consequently real GDP, for the first time since 1993. Other African economies have been following suit. Nigeria, already one of Africa’s largest economies, added 89 percent to its GDP in one swoop in 2014, taking it well past South Africa in size, just by taking account of changes including the growth of booming industries like mobile telecommunications and Nollywood movies.<sup>26</sup> Kenya, another country with

a rapidly growing mobile technology sector, added 25 per cent to its level of GDP through statistical revisions. Other African countries will be following suit. So Africa as a whole probably is not as poor as we've long thought, although nothing real in any of these economies has changed with the revisions, and they are certainly nowhere near as rich as the United Kingdom or the United States.

The trouble with using old weights is that the structure of the economy changes quite dramatically over time. In many African, Asian, and Latin American economies, the GDP calculations take no account of phenomena such as globalization, or the mobile phone revolution in the developing world. The African governments have been keen to update their statistics recently, to tell a different story about their countries, but China's has not. Its data revisions in 2014 added 3.4 per cent to the level of GDP but China did not—for now—adopt the international 'System of National Accounts,' which would have led to a much larger upward revision, because its government prefers China not to be seen as a rich country.

Donors to the poor countries have been funding efforts to improve the way real GDP is calculated—there is an initiative known as PARIS21 (Partnership in Statistics for Development in the 21st Century)—but the planned improvements stretch out to 2020 and beyond.<sup>27</sup> There are fundamental weaknesses with the collection of basic statistics such as what businesses there are, what they are selling, or what goods and services households spend their incomes on. The surveys needed to collect this information are carried out only infrequently. In fact, a recent study found that in the data set frequently used by economists to make international comparisons, twenty-four out of forty-five countries had no price survey data at all.<sup>28</sup> Some countries are using weights that have not been changed since 1968, and only ten sub-Saharan

African countries use weights less than a decade old.<sup>29</sup> In each case where old weights have been used for years, there will be large upward revisions in estimated real GDP when the weights are updated. This could profoundly change our impression of the character and weakness or strength of these economies; one estimate suggests that for twenty years sub-Saharan African economies have been growing three times faster than suggested by the “official” data (although, to emphasize the point, there has been no change in the extent of poverty or people’s living standards because of the statistical changes).<sup>30</sup>

For this reason, the developed economies’ national accounts for the most part now use a “chain-weighted” price index in the calculation of real GDP, meaning that the weights used to combine the separate prices into one index change steadily year by year. Otherwise, as just described, the weights in any given base year would diverge further and further from the actual pattern of the economy. The main catch with this method is that the “real terms” components of GDP no longer add up to the total: the equation  $C + I + G + (X - M) = \text{GDP}$  no longer holds for the chain-weight inflation-adjusted figures. (This is because there is always a residual left over, often small but not always, when some prices and therefore weights have changed a lot.)

Chain-weighting also tells a different big-picture story about the economy, just as rebasing does. For example, historic GDP statistics such as those developed by Angus Maddison for the Organization for Economic Cooperation and Development (OECD) have not been recalculated using chain weights. To do so would change the accepted picture of international growth patterns. Maddison noted, “Acceptance of the new measure for this period [pre-1950] would involve a major reinterpretation of American history.” It would show

U.S. productivity lower than the United Kingdom's in 1914, for example, and much lower U.S. growth and level of GDP than the United Kingdom's by 1929.<sup>31</sup> This is certainly not the received wisdom among economic historians, which matters because their explanations of what drives growth—with great relevance for policies now—could turn out to be based on an inaccurate understanding of what the economy was “really” doing in the nineteenth and twentieth centuries. Just as with developing countries now, changing the method of calculating a price index presents a different pattern of growth. The seemingly technical issue of how best to calculate a price index has some profound implications, ones which economists routinely ignore. In short, the choice of techniques completely alters even the broad outlines of the big picture on economic growth.

### *Silicon Valley Statistical Headaches*

Silicon Valley has caused official statisticians no end of headaches. The price you had to pay for a laptop may have declined somewhat, but the price you had to pay for a unit of processing power absolutely plummeted. Similarly with other products and services such as cameras or mobile telephones or Internet connectivity. This is another particularly thorny problem related to inflation in the prices of certain types of item. Sometimes an increase in the price of an item reflects an improvement in *quality*, and failing to take account of this would lead to an underestimate of real GDP. For many years, this issue was ignored. But the amazingly rapid improvement in computers and consumer electronics from the mid-1990s made it too big to sweep under the carpet. In the United States this question was investigated by the Boskin Commission. Its 1996 report concluded that failure to take account

of productivity improvements had led to an overstatement of inflation (and corresponding understatement of real growth) of 1.3 percentage points a year before 1996, about half of this due to new products and quality improvements.<sup>32</sup> Subsequently, official statisticians in most countries have gradually switched to using a method called “hedonic price” measurement to calculate a price index for goods and services of this kind. This involves taking the actual prices paid and calculating how those relate to various quality characteristics of the items in question. It is yet another complexity in the production of GDP statistics. The final chapter will pick up on this issue again.

### *Britain's Noncrisis of 1976*

The 1970s, as discussed in chapter 3, were a terrible decade for the economy (as well as for fashion and haircuts). Growth was low; inflation was high. Britain did particularly badly and had a ballooning trade deficit—so large, relative to GDP, that it was not clear the country would have enough foreign currency to pay for its imports. The financial markets lost confidence and the value of the pound plunged. The chancellor of the exchequer Denis Healey was on his way to the airport for a trip to Washington when, dramatically, he had to turn back and give a press conference announcing that the United Kingdom was asking the IMF for an emergency loan. Its condition was that the government's fiscal deficit had to be slashed as a proportion of GDP. The Labour government introduced savage public spending cuts. Three years later, Margaret Thatcher swept to power in a Conservative government. Later—some time later—both the borrowing and the GDP figures were revised to an extent suggesting that the “crisis”

hadn't actually been all that bad. Reflecting on the crisis in later years, Healey said, "If we had had the right figures, we would never have needed to go for the loan."<sup>33</sup> Who knows whether Mrs. Thatcher would have won the same kind of election victory if her predecessors in power had not had to bring in the IMF?

There is always pressure on the statisticians to produce timely data, so early estimates of the previous quarter's GDP will almost always be revised subsequently as more of the component sources of data become available. These revisions can be significant, which makes life frustrating for policy-makers trying to figure out what, if anything, they need to do to respond to the business cycle. Although the notion of "fine-tuning" the economy by adjusting tax and spending or interest rates has been pretty much discredited among economists, thanks to the terrible experience of this approach going wrong in the 1970s, there is still immense pressure on politicians and central bankers to try to boost the growth of GDP during a recession. The years of lackluster growth seen since the start of the financial crisis in 2008 are just such a period; a 2014 upward revision of 0.1 percentage point to past data removed the 'double dip' recession everybody thought had occurred post-crisis. Even though a quarterly decline of, say, 0.1 or 0.2 percent in GDP could easily be revised to zero or an increase as the weeks go by and new estimates are published, knowing this seems to do nothing to reduce the expectation that the government should take action.

This isn't the end of the list of practical but consequential difficulties in constructing GDP figures. There is the question of actually *collecting* all the statistics. Many, many separate sources of data are used, from large-scale surveys of the economy such as the five-yearly economic census by the U.S.

Bureau of Economic Analysis, to monthly figures for the production of specific goods collected by trade bodies or in replies to official questionnaires sent to a sample of businesses, to prices sampled by the statistical bureaus, to tax returns, and many more. There are, once again, numerous practical issues. One is the constant difficulty of collecting data measuring the services sector of the economy, now the major part of GDP. The standard surveys used to collect information from businesses do not cover much of the service sector. Another example is the difficulty of keeping track of changing purchasing habits. Consumers have moved progressively from purchasing in local stores to shopping in large stores, including discount “big-box” stores—where businesses may also buy some of their supplies. Now spending is shifting online. A third example is estimating the value of income received in the form of deferred stock options, once a small part of total remuneration but now quite significant. Anyway, there are errors in collecting the figures; I met one man who told me it used to be his job to fill out the forms from the statistics office, but he was too busy so he simply added a bit to the previous year’s numbers.

The actual number for GDP is therefore the product of a vast patchwork of statistics and a complicated set of processes carried out on the raw data to fit them to the conceptual framework.

### *The “Production Boundary”*

As if all of this were not enough, there are some important conceptual questions about the GDP definition, some of which will be followed up in later chapters. The definitions have evolved over the years, and there are areas of active debate among national statistics experts.

Much of GDP is private-sector output or expenditure measured at the prices charged in the market, as mentioned earlier. But large portions of output are not marketed—everything done by the government, for one thing. This has to be valued in a variety of other ways, such as the wages paid to government employees. Some of the government spending ought to be deducted, as it consists of intermediate goods: just as the purchases of nails by a furniture manufacturer would be removed to avoid double counting, so should the part of government spending such as waste removal or fire-fighting services that are intermediate inputs to final output. But that isn't done, mainly because the allocation between final and intermediate public services is impossible in practice.

Other nonmarketed items include the value homeowners get from their own homes despite not having to pay rent; the statisticians “impute” a value to this by looking at rents in the market. But some nonmarketed parts of output, such as unpaid work in the home, are not counted on the grounds that this is also too hard to measure. This results in the paradox (to which we will return in chapter 5) that a widower who marries his housekeeper and stops paying her a wage reduces GDP.

Decisions like these concern what is known as the “production boundary”: what counts as economic output? Government expenditure and household services (such as cleaning and home-grown vegetables) are just two obvious fuzzy areas. The OECD's *Handbook* comments: “There is general consensus favouring the inclusion in GDP of the services provided by general government. Although these services are not sold, they are included as output (value added) in the national accounts and called non-market services produced by general government. This value added is very substantial, since it represents roughly 15% to 20% of GDP, depending

on the OECD country concerned.”<sup>34</sup> As we will see, however, this “general consensus” is relatively recent.

There is an equally general consensus that household services should not be included—although including their value (as discussed in chapter 6) might add 50 percent or more to estimates of GDP, much more than the scale of the government contribution.

There is a wider issue about production on “own-account,” as it’s described. Just as households decide, in producing meals, whether to grow vegetables or buy them, businesses have to decide whether to buy supplies, such as components or payroll services, or produce them for themselves. If they choose production for themselves, these components and services will not be measured in GDP; they will be used up during the process of production. If they outsource, however, these components and services will count toward the output measured in GDP. For this reason, national statisticians focus on the idea of value added, a measure independent of the way firms organize themselves. It is the value a firm adds to the intermediate goods and services it uses. Value added is equal to output (sales plus change in inventories) less purchases.

Besides, it isn’t always clear when to count purchases as intermediate goods or, by contrast, investment spending. Before 2008, national accounts treated businesses’ spending on R&D in the same way as spending on raw materials or cleaners; that is, as an intermediate good that would not count toward final output. This is starting to change; henceforth, R&D is supposed to be counted as investment. This change of category has already been applied in theory (since 1993) to spending on software, leading to upward revisions of 1–4 percent in the level of GDP, depending on the country, but in practice it is hard to implement because many firms do not record their spending on software as an investment.

There is a final area of the economy that poses great difficulty, and that is financial services—chapter 5 will return to this in more depth. The financial crisis has clearly raised some questions about whether the statistical treatment of financial services is right: are they making the contribution to the economy we were all assuming before 2008?

Finally, and above all, it is important to recognize that GDP is **not** a measure of welfare. It—notoriously—counts the services of lawyers and other “bads” as a positive. So too the rebuilding of bridges and homes after a storm like Katrina or Sandy, or floods. GDP measures output; it does not measure well-being. Much more on this, too, as this brief history unfolds.

