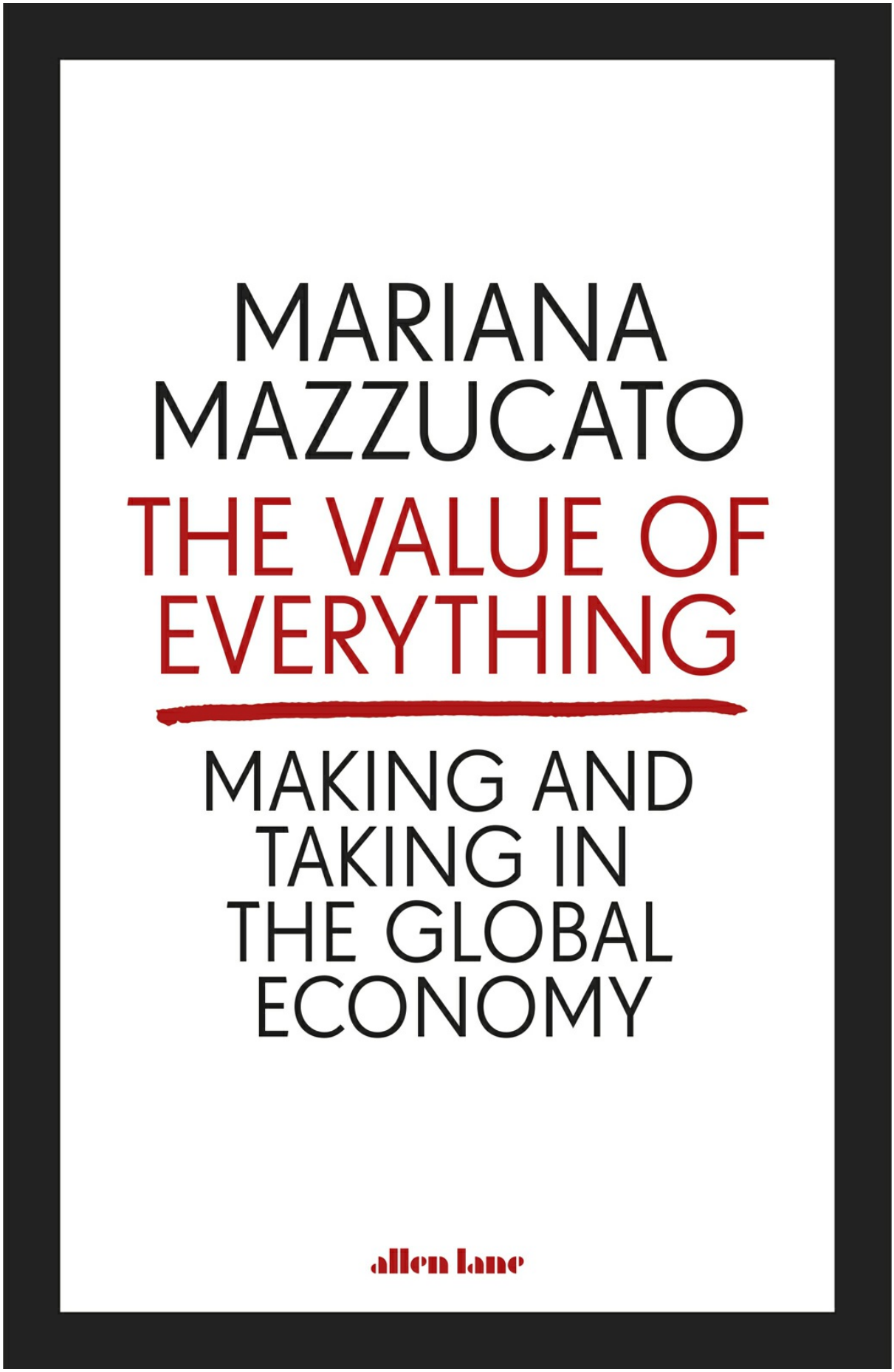


Mariana Mazzucato

THE VALUE OF EVERYTHING

Making and Taking in the Global Economy

2018

Contents

Preface: Stories About Wealth Creation

We often hear businesses, entrepreneurs or sectors talking about themselves as ‘wealth-creating’. The contexts may differ – finance, big pharma or small start-ups – but the self-descriptions are similar: I am a particularly productive member of the economy, my activities create wealth, I take big ‘risks’, and so I deserve a higher income than people who simply benefit from the spillovers of this activity. But what if, in the end, these descriptions are simply just stories? Narratives created in order to justify inequalities of wealth and income, massively rewarding the few who are able to convince governments and society that they deserve high rewards, while the rest of us make do with the leftovers.

In 2009 Lloyd Blankfein, CEO of Goldman Sachs, claimed that ‘The people of Goldman Sachs are among the most productive in the world.’[1] Yet, just the year before, Goldman had been a major contributor to the worst financial and economic crisis since the 1930s. US taxpayers had to stump up $125 billion to bail it out. In light of the terrible performance of the investment bank just a year before, such a bullish statement by the CEO was extraordinary. The bank laid off 3,000 employees between November 2007 and December 2009, and profits plunged.[2] The bank and some its competitors were fined, although the amounts seemed small relative to later profits: fines of $550 million for Goldman and $297 million for J. P. Morgan, for example.[3] Despite everything, Goldman – along with other banks and hedge funds – proceeded to bet against the very instruments which they had created and which had led to such turmoil.

Although there was much talk about punishing those banks that had contributed to the crisis, no banker was jailed, and the changes hardly dented the banks’ ability to continue making money from speculation: between 2009 and 2016 Goldman achieved net earnings of $63 billion on net revenues of $250 billion.[4] In 2009 alone they had record earnings of $13.4 billion.[5] And although the US government saved the banking system with taxpayers’ money, the government did not have the confidence to

demand a fee from the banks for such high-risk activity. It was simply happy, in the end, to get its money back.

Financial crises, of course, are not new. Yet Blankfein's exuberant confidence in his bank would have been less common half a century ago. Until the 1960s, finance was not widely considered a 'productive' part of the economy. It was viewed as important for transferring existing wealth, not creating new wealth. Indeed, economists were so convinced about the purely facilitating role of finance that they did not even include most of the services that banks performed, such as taking in deposits and giving out loans, in their calculations of how many goods and services are produced by the economy. Finance sneaked into their measurements of Gross Domestic Product (GDP) only as an 'intermediate input' – a service contributing to the functioning of other industries that were the real value creators.

In around 1970, however, things started to change. The national accounts – which provide a statistical picture of the size, composition and direction of an economy – began to include the financial sector in their calculations of GDP, the total value of the goods and services produced by the economy in question.[6] This change in accounting coincided with the deregulation of the financial sector which, among other things, relaxed controls on how much banks could lend, the interest rates they could charge and the products they could sell. Together, these changes fundamentally altered how the financial sector behaved, and increased its influence on the 'real' economy. No longer was finance seen as a staid career. Instead, it became a fast track for smart people to make a great deal of money. Indeed, after the Berlin Wall fell in 1989, some of the cleverest scientists in Eastern Europe ended up going to work for Wall Street. The industry expanded, grew more confident. It openly lobbied to advance its interests, claiming that finance was critical for wealth creation.

Today the issue is not just the size of the financial sector, and how it has outpaced the growth of the non-financial economy (e.g. industry), but its effect on the behaviour of the rest of the economy, large parts of which have been 'financialized'. Financial operations and the mentality they breed pervade industry, as can be seen when managers choose to spend a greater proportion of profits on share buy-backs – which in turn boost stock prices, stock options and the pay of top executives – than on investing in the long-term future of the business. They call it value creation but, as in the financial sector itself, the reality is often the opposite: value extraction.

These stories of value creation are not limited to finance. In 2014 the pharmaceutical giant Gilead priced its new treatment for the life-threatening hepatitis C virus, Harvoni, at $94,500 for a three-month course. Gilead justified charging this price by insisting that it represented 'value' to health systems. John LaMattina, former President of R&D at the drugs company Pfizer, argued that the high price of speciality drugs is justified by how beneficial they are for patients and for society in general. In practice, this means relating the price of a drug to the costs that the disease would cause to society if not treated, or if treated with the second-best therapy available. The industry calls this 'value-based pricing'. It's an argument refuted by critics, who cite case studies that show no correlation between the price of cancer drugs and the benefits they provide.[7] One interactive calculator (www.drugabacus.org), which enables you to establish the 'correct' price of a cancer drug on the basis of its valuable characteristics (the increase in life expectancy it provides to patients, its side effects, and so on), shows that for most drugs this value-based price is lower than the current market price.[8]

Yet drug prices are not falling. It seems that the industry's value creation arguments have successfully neutralized criticism. Indeed, a high proportion of health care costs in the Western world has nothing to do with health care: these costs are simply the value the pharmaceutical industry extracts.

Or consider the way that entrepreneurs in the dot.com and IT industry lobby for advantageous tax treatment by governments in the name of 'wealth creation'. With 'innovation' as the new force in modern capitalism, Silicon Valley's do-gooders have successfully projected themselves as the entrepreneurs and garage tinkerers who unleash the 'creative destruction' from which the jobs of the future come. These new actors, from Google to Uber to Airbnb, are often described as the 'wealth creators'.

Yet this seductive story of value creation leads to questionable broader tax policies by policymakers: for example, the 'patent box' policy that reduces tax on any products whose inputs are patented, supposedly to incentivize innovation by rewarding the generation of intellectual property. It's a policy that makes little sense, as patents are already monopolies which should normally earn high returns. Policymakers' objectives should not be to increase the profits from monopolies, but to favour investments in areas like research.

Many of the so-called wealth creators in the tech industry, like the co-founder of Pay Pal, Peter Thiel, often lambast government as a pure

impediment to wealth creation.[9] Thiel went so far as to set up a ‘secessionist movement’ in California so that the wealth creators could be as independent as possible from the heavy hand of government. When Eric Schmidt, CEO of Google, was quizzed about the way companies control our personal data, he replied with what he assumed was a rhetorical question: ‘Would you prefer government to have it?’ His reply fed a modern-day banality: entrepreneurs good, government bad.

Yet in presenting themselves as modern-day heroes, Apple and other companies conveniently ignore the pioneering role of government in new technologies. Apple has unashamedly declared that its contribution to society should not be sought through tax but through recognition of its great gizmos. But doesn’t the taxpayer who helped Apple create those products and the record profits and cash mountain they have generated deserve something back, beyond a series of undoubtedly brilliant gadgets? Simply to pose this question, however, underlines how we need a radically different type of narrative as to who created the wealth in the first place – and who has subsequently extracted it.

If there are so many wealth creators in industry, the inevitable conclusion is that at the opposite side of the spectrum featuring fleet-footed bankers, science-based pharmaceuticals and entrepreneurial geeks are the inert, value-extracting civil servants and bureaucrats in government. In this view, if private enterprise is the fast cheetah bringing innovation to the world, government is a plodding tortoise impeding progress – or, to invoke a different metaphor, a Kafkaesque bureaucrat, buried under papers, cumbersome and inefficient. Government is depicted as a drain on society, funded by obligatory taxes on long-suffering citizens. In this story, there is always only one conclusion: that we need more market and less state. The slimmer, trimmer and more efficient the state machine the better.

In all these cases, from finance to pharmaceuticals and IT, governments bend over backwards to attract these supposedly value-creating individuals and companies, dangling before them tax reductions and exemptions from the red tape that is believed to constrict their wealth-creating energies. The media heap wealth creators with praise, politicians court them, and for many people they are high-status figures to be admired and emulated. But who decided that they are creating value? What definition of value is used to distinguish value creation from value extraction, or even from value destruction?

Why have we so readily believed this narrative of good versus bad? How is the value produced by the public sector measured, and why is it

more often than not treated simply as a more inefficient version of the private sector? What if there was actually no evidence for this story at all? What if it stemmed purely from a set of deeply ingrained ideas? What new stories might we tell?

The Greek philosopher Plato once argued that storytellers rule the world. His great work *The Republic* is in part a guide to educating the leader of his ideal state, the Guardian. Plato recognized that stories form character, culture and behaviour: 'Our first business is to supervise the production of stories, and chose only those we think suitable, and reject the rest. We shall persuade mothers and nurses to tell our chosen stories to their children, and by means of them to mould their minds and characters rather than their bodies. The greater part of the stories current today we shall have to reject.'[10]

Plato disliked myths about ill-behaved gods. This book looks at a more modern myth, about value creation in the economy. Such myth-making, I argue, has allowed an immense amount of value extraction, enabling some individuals to become very rich and draining societal wealth in the process. Between 1975 and 2015 real US GDP – the size of the economy adjusted for inflation – roughly tripled from $5.49 trillion to $16.58 trillion.[11] During this period, productivity grew by more than 60 per cent. Yet from 1979 onwards, real hourly wages for the great majority of American workers have stagnated or even fallen.[12] In other words, for almost four decades a tiny elite has captured nearly all the gains from an expanding economy. You do not have to look far to see who is in that elite. Mark Zuckerberg dropped out of Harvard at the age of nineteen to launch Facebook. He is now in his early thirties. According to Forbes,[13] Zuckerberg's fortune increased by $18 billion in the year to mid-2016, making his current total estimated worth $70.8 billion. He is the fourth-richest man in the US and the fifth-richest in the world.[14]

It defies reason to maintain, as the dominant narrative does, that the inequality that has increased in the US, and in many other economies, is due to very smart individuals doing particularly well in innovative sectors. While wealth is created through a collective effort, the massive imbalance in the distribution of the gains from economic growth has often been more the result of wealth extraction, whose potential scale globalization has greatly magnified.

At the end of the second quarter of 2016, Facebook had 1.71 billion monthly active users, almost one in every four people on the planet. The imbalance in the distribution of gains from economic growth is a primary cause of widening social inequalities in many mature economies, which in

turn has deep political consequences – arguably including the UK's referendum vote to leave the European Union. Many people who felt globalization had left them behind chose Brexit.

Economists must take a sizeable share of the blame for the lamentable outcomes of the prevailing story about value. We have stopped debating value – and, as a result, we have allowed one story about 'wealth creation' and 'wealth creators' to dominate almost unchallenged.

The purpose of this book is to change this state of things, and to do so by reinvigorating the debate about value that used to be – and, I argue, should still be – at the core of economic thinking. If value is defined by price – set by the supposed forces of supply and demand – then as long as an activity fetches a price (legally), it is seen as creating value. So if you earn a lot you must be a value creator. I will argue that the way the word 'value' is used in modern economics has made it easier for value-extracting activities to masquerade as value-creating activities. And in the process rents (unearned income) get confused with profits (earned income); inequality rises, and investment in the real economy falls. What's more, if we cannot differentiate value creation from value extraction, it becomes nearly impossible to reward the former over the latter. If the goal is to produce growth that is more innovation-led (smart growth), more inclusive and more sustainable, we need a better understanding of value to steer us.

This is not an abstract debate. It has far-reaching consequences – social and political as well as economic – for everyone. How we discuss value affects the way all of us, from giant corporations to the most modest shopper, behave as actors in the economy and in turn feeds back into the economy, and how we measure its performance. This is what philosophers call 'performativity': how we talk about things affects behaviour, and in turn how we theorize things. In other words, it is a self-fulfilling prophecy.

If we cannot define what we mean by value, we cannot be sure to produce it, nor to share it fairly, nor to sustain economic growth. The understanding of value, then, is critical to all the other conversations we need to have about where our economy is going and how to change its course.

Introduction: Making versus Taking

> *The barbarous gold barons – they did not find the gold, they did not mine the gold, they did not mill the gold, but by some weird alchemy all the gold belonged to them.*
>
> Big Bill Haywood, founder of the Unites States' first industrial union[1]

Bill Haywood expressed his puzzlement eloquently. He represented men and women in the US mining industry at the start of the twentieth century and during the Great Depression of the 1930s. He was steeped in the industry. But even Haywood could not answer the question: why did the owners of capital, who did little but buy and sell gold on the market, make so much money, while workers who expended their mental and physical energy to find it, mine it and mill it, make so little? Why were the *takers* making so much money at the expense of the *makers*?

Similar questions are still being asked today. In 2016 the British high-street retailer BHS collapsed. It had been founded in 1928 and in 2004 was bought by Sir Philip Green, a well-known retail entrepreneur, for £200 million. In 2015 Sir Philip sold the business for £1 to a group of investors headed by the British businessman Dominic Chappell. While it was under his control, Sir Philip and his family extracted from BHS an estimated £580 million in dividends, rental payments and interest on loans they had made to the company. The collapse of BHS threw 11,000 people out of work and left its pension fund with a £571 million deficit, even though the fund had been in surplus when Sir Philip acquired it.[2] A report on the BHS disaster by the House of Commons Work and Pensions Select Committee accused Sir Philip, Mr Chappell and their 'hangers-on' of 'systematic plunder'. For BHS workers and pensioners who depended on the company for a decent living for their families, this was value extraction – the appropriation of gains vastly out of proportion to economic contribution – on an epic scale. For Sir Philip and others who controlled the business, it was value creation.

While Sir Philip's activities could be viewed as an aberration, the excesses of an individual, his way of thinking is hardly unusual: today, many giant corporations are also guilty of confusing value creation with value extraction. In August 2016, for instance, the European Commission,

the European Union's (EU) executive arm, sparked an international row between the EU and the US when it ordered Apple to pay €13 billion in back taxes to Ireland.[3]

Apple is the world's biggest company by stock market value. In 2015 it held a mountain of cash and securities outside the US worth $187 billion[4] – about the same size as the Czech Republic's economy that year[5] – to avoid paying the US taxes that would be due on the profits if they were repatriated. Under a deal with Ireland dating back to 1991, two Irish subsidiaries of Apple received very generous tax treatment. The subsidiaries were Apple Sales International (ASI), which recorded all the profits earned on sales of iPhones and other Apple devices in Europe, the Middle East, Africa and India; and Apple Operations Europe, which made computers. Apple transferred development rights of its products to ASI for a nominal amount, thereby depriving the US taxpayer of revenues from technologies, embodied in Apple products, whose early development the taxpayer had funded. The European Commission alleged that the maximum rate payable on those profits booked through Ireland which were liable for tax was 1 per cent, but that in 2014 Apple paid tax at 0.005 per cent. The usual rate of corporation tax in Ireland is 12.5 per cent.

What is more, these 'Irish' subsidiaries of Apple are in fact not resident for tax purposes anywhere. This is because they have exploited discrepancies between the Irish and US definitions of residence. Almost all the profits earned by the subsidiaries were allocated to their 'head offices', which existed only on paper. The Commission ordered Apple to pay the back taxes on the grounds that Ireland's deal with Apple constituted illegal state aid (government support that gives a company an advantage over its competitors); Ireland had not offered other companies similar terms. Ireland, the Commission alleged, had offered Apple ultra-low taxes in return for the creation of jobs in other Apple businesses there. Apple and Ireland rejected the Commission's demand – and of course Apple is not the only major corporation to have constructed exotic tax structures.

But Apple's value extraction cycle is not limited to its international tax operations – it is also much closer to home. Not only did Apple extract value from Irish taxpayers, but the Irish government has extracted value from the US taxpayer. Why so? Apple created its intellectual property in California, where its headquarters are based. Indeed, as I argued in my previous book, *The Entrepreneurial State*,[6] and discuss briefly here in Chapter 7, all the technology that makes the smartphone smart was publicly funded. But in 2006 Apple formed a subsidiary in Reno, Nevada, where there is no corporate income or capital gains tax, in order to avoid

state taxes in California. Creatively naming it Braeburn Capital, Apple channelled a portion of its US profits to the Nevada subsidiary instead of reporting it in California. Between 2006 and 2012, Apple earned $2.5 billion in interest and dividends reported in Nevada to avoid Californian tax. California's infamously large debt would be significantly reduced if Apple fully and accurately reported its US revenues in that state, where a major portion of its value (architecture, design, sales, marketing and so on) originated. Value extraction thus pits US states against each other, as well as the US against other countries.

It is clear that Apple's highly complex tax arrangements were principally designed to extract the maximum value from its business by avoiding paying substantial taxes which would have benefited the societies in which the company operated. Apple certainly creates value, of that there is no doubt: but to ignore the support taxpayers have given it, and then to pit states and countries against each other, is surely not the way to build an innovative economy or achieve growth that is inclusive, that benefits a wide section of the population, not only those best able to 'game' the system.

There is yet another dimension to Apple's value extraction. Many such corporations use their profits to boost share prices in the short term instead of reinvesting them in production for the long term. The main way they do this is by using cash reserves to buy back shares from investors, arguing that this is to maximize shareholder 'value' (the income earned by shareholders in the company, based on the valuation of the company's stock price). But it is no accident that among the primary beneficiaries of share buy-backs are managers with generous share option schemes as part of their remuneration packages – the same managers who implement the share buy-back programmes. In 2012, for example, Apple announced a share buy-back programme of up to a staggering $100 billion, partly to ward off 'activist' shareholders demanding that the company return cash to them to 'unlock shareholder value'.[7] Rather than reinvest in the business, Apple preferred to transfer cash to shareholders.

The alchemy of the takers versus the makers that Big Bill Haywood referred to back in the 1920s continues today.

COMMON CRITIQUES OF VALUE EXTRACTION

The vital but often muddled distinction between value extraction and value creation has consequences far beyond the fate of companies and their workers, or even of whole societies. The social, economic and political impacts of value extraction are huge. Prior to the 2007 financial crisis, the

income share of the top 1 per cent in the US expanded from 9.4 per cent in 1980 to a staggering 22.6 per cent in 2007. And things are only getting worse. Since 2009 inequality has been increasing even more rapidly than before the 2008 financial crash. In 2015 the combined wealth of the planet's sixty-two richest individuals was estimated to be about the same as that of the bottom half of the world's population – 3.5 billion people.[8]

So how does the alchemy continue to happen? A common critique of contemporary capitalism is that it rewards 'rent seekers' over true 'wealth creators'. 'Rent-seeking' here refers to the attempt to generate income, not by producing anything new but by overcharging above the 'competitive price', and undercutting competition by exploiting particular advantages (including labour), or, in the case of an industry with large firms, their ability to block other companies from entering that industry, thereby retaining a monopoly advantage. Rent-seeking activity is often described in other ways: the 'takers' winning out over the 'makers', and 'predatory' capitalism winning over 'productive' capitalism. It's seen as a key way – perhaps *the* key way – in which the 1 per cent have risen to power over the 99 per cent.[9] The usual targets of such criticism are the banks and other financial institutions. They are seen as profiting from speculative activities based on little more than buying low and selling high, or buying and then stripping productive assets simply to sell them on again with no real value added.

More sophisticated analyses have linked rising inequality to the particular way in which the 'takers' have increased their wealth. The French economist Thomas Piketty's influential book *Capital in the Twenty-First Century* focuses on the inequality created by a predatory financial industry that is taxed insufficiently, and by ways in which wealth is inherited across generations, which gives the richest a head start in getting even richer. Piketty's analysis is key to understanding why the rate of return on financial assets (which he calls capital) has been higher than that on growth, and calls for higher taxes on the resultant wealth and inheritance to stop the vicious circle. Ideally, from his point of view, taxes of this sort should be global, so as to avoid one country undercutting another.

Another leading thinker, the US economist Joseph Stiglitz, has explored how weak regulation and monopolistic practices have allowed what economists call 'rent extraction', which he sees as the main impetus behind the rise of the 1 per cent in the US.[10] For Stiglitz, this rent is the income obtained by creating impediments to other businesses, such as barriers to prevent new companies from entering a sector, or deregulation

that has allowed finance to become disproportionately large in relation to the rest of the economy. The assumption is that, with fewer impediments to the functioning of economic competition, there will be a more equal distribution of income.[11]

I think we can go even further with these 'makers' versus 'takers' analyses of why our economy, with its glaring inequalities of income and wealth, has gone so wrong. To understand how some are perceived as 'extracting value', siphoning wealth away from national economies, while others are 'wealth creators' but do not benefit from that wealth, it is not enough to look at impediments to an idealized form of perfect competition. Yet mainstream ideas about rent do not fundamentally challenge how value extraction occurs – which is why it persists.

In order to tackle these issues at root, we need to examine where *value* comes from in the first place. What exactly is it that is being extracted? What social, economic and organizational conditions are needed for value to be produced? Even Stiglitz's and Piketty's use of the term 'rent' to analyse inequality will be influenced by their idea of what value is and what it represents. Is rent simply an impediment to 'free-market' exchange? Or is it due to their positions of power that some can earn 'unearned income' – that is, income derived from moving existing assets around rather than creating new ones?[12] This is a key question we will look at in Chapter 2.

WHAT IS VALUE?

Value can be defined in different ways, but at its heart it is the production of new goods and services. How these outputs are produced (production), how they are shared across the economy (distribution) and what is done with the earnings that are created from their production (reinvestment) are key questions in defining economic value. Also crucial is whether what it is that is being created is useful: are the products and services being created increasing or decreasing the resilience of the productive system? For example, it might be that a new factory is produced that is valuable economically, but if it pollutes so much to destroy the system around it, it could be seen as not valuable.

By 'value creation' I mean the ways in which different types of resources (human, physical and intangible) are established and interact to produce new goods and services. By 'value extraction' I mean activities focused on moving around existing resources and outputs, and gaining disproportionately from the ensuing trade.

A note of caution is important. In the book I use the words ‘wealth’ and ‘value’ almost interchangeably. Some might argue against this, seeing wealth as a more monetary and value as potentially a more social concept, involving not only *value* but *values*. I want to be clear on how these two words are used. I use ‘value’ in terms of the ‘process’ by which wealth is created – it is a flow. This flow of course results in actual things, whether tangible (a loaf of bread) or intangible (new knowledge). ‘Wealth’ instead is regarded as a cumulative stock of the value already created. The book focuses on value and what forces produce it – the process. But it also looks at the claims around this process, which are often phrased in terms of ‘who’ the wealth creators are. In this sense the words are used interchangeably.

For a long time the idea of value was at the heart of debates about the economy, production and the distribution of the resulting income, and there were healthy disagreements over what value actually resided in. For some economic schools of thought, the price of products resulted from supply and demand, but the value of those products derived from the amount of work that was needed to produce things, the ways in which technological and organizational changes were affecting work, and the relations between capital and labour. Later, this emphasis on ‘objective’ conditions of production, technology and power relationships was replaced by concepts of scarcity and the ‘preferences’ of economic actors: the amount of work supplied is determined by workers’ preference for leisure over earning a higher amount of money. Value, in other words, became *subjective*.

Until the mid-nineteenth century, too, almost all economists assumed that in order to understand the prices of goods and services it was first necessary to have an objective theory of value, a theory tied to the conditions in which those goods and services were produced, including the time needed to produce them, the quality of the labour employed; and the determinants of ‘value’ actually shaped the price of goods and services. Then, this thinking began to go into reverse. Many economists came to believe that the value of things was determined by the price paid on the ‘market’ – or, in other words, what the consumer was prepared to pay. All of a sudden, value was in the eye of the beholder. Any goods or services being sold at an agreed market price were by definition value-creating.

The swing from value determining price to price determining value coincided with major social changes at the end of the nineteenth century. One was the rise of socialism, which partly based its demands for reforms on the claim that labour was not being rewarded fairly for the value it

created, and the ensuing consolidation of a capitalist class of producers. The latter group was, unsurprisingly, keen on the alternative theory, that price determined value, a story which allowed them to defend their appropriation of a larger share of output, with labour increasingly being left behind.

In the intellectual world, economists wanted to make their discipline seem 'scientific' – more like physics and less like sociology – with the result that they dispensed with its earlier political and social connotations. While Adam Smith's writings were full of politics and philosophy, as well as early thinking about how the economy works, by the early twentieth century the field which for 200 years had been 'political economy' emerged cleansed as simply 'economics'. And economics told a very different story.

Eventually the debate about different theories of value and the dynamics of value creation virtually vanished from economics departments, only showing up in business schools in a very new form: 'shareholder value',[13] 'shared value',[14] 'value chains',[15] 'value for money', 'valuation', 'adding value' and the like. So while economics students used to get a rich and varied education in the idea of value, learning what different schools of economic thought had to say about it, today they are taught only that value is determined by the dynamics of price, due to scarcity and preferences. This is not presented as a particular theory of value – just as Economics 101, the introduction to the subject. An intellectually impoverished idea of value is just taken as read, assumed simply to be true. And the disappearance of the concept of value, this book argues, has paradoxically made it much easier for this crucial term 'value' – a concept that lies at the heart of economic thought – to be used and abused in whatever way one might find useful.

MEET THE PRODUCTION BOUNDARY

To understand how different theories of value have evolved over the centuries, it is useful to consider why and how some activities in the economy have been called 'productive' and some 'unproductive', and how this distinction has influenced ideas about which economic actors deserve what – how the spoils of value creation are distributed.

For centuries, economists and policymakers – people who set a plan for an organization such as government or a business – have divided activities according to whether they produce value or not; that is, whether they are productive or unproductive. This has essentially created a boundary – the fence in Figure 1 below – thereby establishing a conceptual boundary –

sometimes referred to as a ‘production boundary’ – between these activities.[16] Inside the boundary are the wealth creators. Outside are the beneficiaries of that wealth, who benefit either because they can extract it through rent-seeking activities, as in the case of a monopoly, or because wealth created in the productive area is redistributed to them, for example through modern welfare policies. Rents, as understood by the classical economists, were *unearned* income and fell squarely outside the production boundary. Profits were instead the returns earned for productive activity inside the boundary.

Historically, the boundary fence has not been fixed. Its shape and size have shifted with social and economic forces. These changes in the boundary between makers and takers can be seen just as clearly in the past as in the modern era. In the eighteenth century there was an outcry when the physiocrats, an early school of economists, called landlords ‘unproductive’. This was an attack on the ruling class of a mainly rural Europe. The politically explosive question was whether landlords were just abusing their power to extract part of the wealth created by their tenant farmers, or whether their contribution of land was essential to the way in which farmers created value.

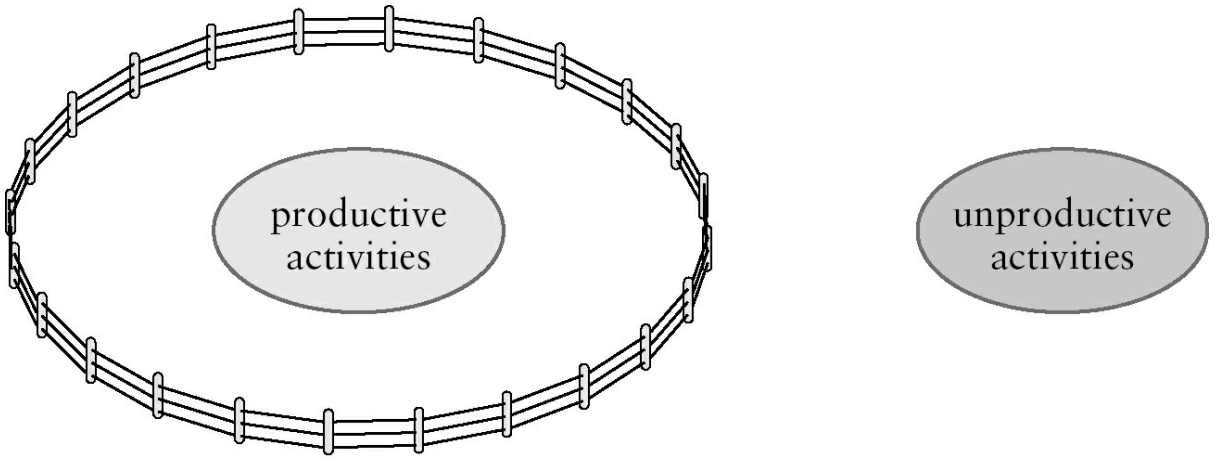


Figure 1. Production boundary around the value-producing activities of the economy

A variation of this debate about where to draw the production boundary continues today with the financial sector. After the 2008 financial crisis, there were calls from many quarters for a revival of industrial policy to boost the ‘makers’ in industry, who were seen to be pitted against the ‘takers’ in finance. It was argued that rebalancing was needed to shrink the size of the financial sector (falling into the dark grey circle of unproductive activities above) by taxation, for example a tax on financial transactions

such as foreign exchange dealing or securities trading, and by industrial policies to nurture growth in industries that actually made things instead of just exchanging them (falling into the light grey circle of productive activities above).

But things are not so simple. The point is not to blame some as takers and to label others as makers. The activities of people outside the boundary may be needed to facilitate production – without their work, productive activities may not be so valuable. Merchants are necessary to ensure the goods arrive at the marketplace and are exchanged efficiently. The financial sector is critical for buyers and sellers to do business with each other. How these activities can be shaped to actually serve their purpose of producing value is the real question.

And, most important of all, what about government? On which side of the production boundary does it lie? Is government inherently unproductive, as is often claimed, its only earnings being compulsory transfers in the form of taxes from the productive part of the economy? If so, how can government make the economy grow? Or can it at best only set the rules of the game, so that the value creators can operate efficiently?

Indeed, the recurring debate about the optimal size of government and the supposed perils of high public debt boils down to whether government spending helps the economy to grow – because government can be productive and add value – or whether it holds back the economy because it is unproductive or even destroys value. The issue is politically loaded and deeply colours current debates, ranging from whether the UK can afford Trident nuclear weapons to whether there is a ‘magic number’ for the size of government, defined as government spending as a proportion of national output, beyond which an economy will inevitably do less well than it might have done if government spending had been lower. As we will explore in Chapter 8, this question is more tainted by political views and ideological positions than informed by deep scientific proofs. Indeed, it is important to remember that economics is at heart a social science, and the ‘natural’ size of government will depend on one’s theory of (or simply ‘position’ on) the purpose of government. If it is seen as useless, or at best a fixer of occasional problems, its optimum size will inevitably be notionally smaller than if it is viewed as a key engine of growth needed to steer and invest in the value creation process.

Over time, this conceptual production boundary was expanded to encompass much more of the economy than before, and more varied economic activities. As economists, and wider society, came to determine value by supply and demand – what is bought has value – activities such as

financial transactions were redefined as productive, whereas previously they had usually been classed as unproductive. Significantly, the only major part of the economy which is now considered largely to lie outside the production boundary – and thus to be 'unproductive' – remains government. It is also true that many other services that people provide throughout society go unpaid, such as care given by parents to children or by the healthy to the unwell, and are not well accounted for. Fortunately, issues such as factoring care into the way we measure national output (GDP) are increasingly coming to the fore. But besides adding new concepts to GDP – such as care, or the sustainability of the planet – it is fundamental to understand why we hold the assumptions about value that we do. And this is impossible if value is not scrutinized.

WHY VALUE THEORY MATTERS

First, the disappearance of value from the economic debate hides what should be alive, public and actively contested.[17] If the assumption that value is in the eye of the beholder is not questioned, some activities will be deemed to be value-creating and others will not, simply because someone – usually someone with a vested interest – says so, perhaps more eloquently than others. Activities can hop from one side of the production boundary to the other with a click of the mouse and hardly anyone notices. If bankers, estate agents and bookmakers claim to create value rather than extract it, mainstream economics offers no basis on which to challenge them, even though the public might view their claims with scepticism. Who can gainsay Lloyd Blankfein when he declares that Goldman Sachs employees are among the most productive in the world? Or when pharmaceutical companies argue that the exorbitantly high price of one of their drugs is due to the *value* it produces? Government officials can become convinced (or 'captured') by stories about wealth creation, as was recently evidenced by the US government's approval of a leukemia drug treatment at half a million dollars, precisely using the 'value-based pricing' model pitched by the industry – even when the taxpayer contributed $200 million dollars towards its discovery.[18]

Second, the lack of analysis of value has massive implications for one particular area: the distribution of income between different members of society. When value is determined by price (rather than vice versa), the level and distribution of income seem justified as long as there is a market for the goods and services which, when bought and sold, generate that income. All income, according to this logic, is earned income: gone is any

analysis of activities in terms of whether they are productive or unproductive.

Yet this reasoning is circular, a closed loop. Incomes are justified by the production of something that is of value. But how do we measure value? By whether it earns income. You earn income because you are productive and you are productive because you earn income. So with a wave of a wand, the concept of *unearned income* vanishes. If income means that we are productive, and we deserve income whenever we are productive, how can income possibly be unearned? As we shall see in Chapter 3, this circular reasoning is reflected in how national accounts – which track and measure production and wealth in the economy – are drawn up. In theory, no income may be judged too high, because in a market economy competition prevents anyone from earning more than he or she deserves. In practice, markets are what economists call imperfect, so prices and wages are often set by the powerful and paid by the weak.

In the prevailing view, prices are set by supply and demand, and any deviation from what is considered the competitive price (based on marginal revenues) must be due to some imperfection which, if removed, will produce the correct distribution of income between actors. The possibility that some activities perpetually earn rent because they are perceived as valuable, while actually blocking the creation of value and/or destroying existing value, is hardly discussed.

Indeed, for economists there is no longer any story other than that of the subjective theory of value, with the market driven by supply and demand. Once impediments to competition are removed, the outcome should benefit everyone. How different notions of value might affect the distribution of revenues between workers, public agencies, managers and shareholders at, say, Google, General Electric or BAE Systems, goes unquestioned.

Third, in trying to steer the economy in particular directions, policymakers are – whether they recognize it or not – inevitably influenced by ideas about value. The rate of GDP growth is obviously important in a world where billions of people still live in dire poverty. But some of the most important economic questions today are about how to achieve a particular type of growth. Today, there is a lot of talk about the need to make growth ‘smarter’ (led by investments in innovation), more sustainable (greener) and more inclusive (producing less inequality).[19]

Contrary to the widespread assumption that policy should be directionless, simply removing barriers and focusing on ‘levelling the playing field’ for businesses, an immense amount of policymaking is

needed to reach these particular objectives. Growth will not somehow go in this direction by itself. Different types of policy are needed to tilt the playing field in the direction deemed desirable. This is very different from the usual assumption that policy should be directionless, simply removing barriers so that businesses can get on with smooth production.

Deciding which activities are more important than others is critical in setting a direction for the economy: put simply, those activities thought to be more important in achieving particular objectives have to be increased and less important ones reduced. We already do this. Certain types of tax credits, for, say, R&D, try to stimulate more investment in innovation. We subsidize education and training for students because as a society we want more young people to go to university or enter the workforce with better skills. Behind such policies may be economic models that show how investment in 'human capital' – people's knowledge and capabilities – benefits a country's growth by increasing its productive capacity. Similarly, today's deepening concern that the financial sector in some countries is too large – compared, for example, to manufacturing – might be informed by theories of what kind of economy we want to be living in and the size and role of finance within it.

But the distinction between productive and unproductive activities has rarely been the result of 'scientific' measurement. Rather, ascribing value, or the lack of it, has always involved malleable socio-economic arguments which derive from a particular political perspective – which is sometimes explicit, sometimes not. The definition of value is always as much about politics, and about particular views on how society ought to be constructed, as it is about narrowly defined economics. Measurements are not neutral: they affect behaviour and vice versa (this is the concept of performativity which we encountered in the Preface).

So the point is not to create a stark divide, labelling some activities as productive and categorizing others as unproductive rent-seeking. I believe we must instead be more forthright in linking our understanding of value creation to the way in which activities (whether in the financial sector or the real economy) should be structured, and how this is connected to the distribution of the rewards generated. Only in this way will the current narrative about value creation be subject to greater scrutiny, and statements such as 'I am a wealth creator' measured against credible ideas about where that wealth comes from. A pharmaceutical company's *value-based pricing* might then be scrutinized with a more collective value-creation process in mind, one in which public money funds a large portion of pharmaceutical research – from which that company benefits – in the

highest-risk stage. Similarly, the 20 per cent share that venture capitalists usually get when a high-tech small company goes public on the stock market may be seen as excessive in light of the actual, not mythological, risk they have taken in investing in the company's development. And if an investment bank makes an enormous profit from the exchange rate instability that affects a country, that profit can be seen as what it really is: rent.

In order to arrive at this understanding of value creation, however, we need to go beyond seemingly scientific categorizations of activities and look at the socio-economic and political conflicts that underlie them. Indeed, claims about value creation have always been linked to assertions about the relative productiveness of certain elements of society, often related to fundamental shifts in the underlying economy: from agricultural to industrial, or from a mass-production-based economy to one based on digital technology.

THE STRUCTURE OF THE BOOK

In Chapters 1 and 2 I look at how economists from the seventeenth century onwards have thought about steering growth by increasing productive activities and reducing unproductive ones, something they conceptualized by means of a theoretical production boundary. The production boundary debate, and its close relationship to ideas of value, has influenced government measures of economic growth for centuries; the boundary, too, has changed, influenced by fluctuating social, economic and political conditions. Chapter 2 delves into the biggest shift of all. From the second half of the nineteenth century onwards, value went from being an objective category to a more subjective one tied to individual preferences. The implications of this revolution were seismic. The production boundary itself was blurred, because almost anything that could get a price or could successfully claim to create value – for example, finance – suddenly became productive. This opened the way to increased inequality, driven by particular agents in the economy being able to brag about their extraordinary 'productivity'.

As we will see in Chapter 3, which explores the development of national accounts, the idea of the production boundary continues to influence the concept of output. There is, however, a fundamental distinction between this new boundary and its predecessors. Today, decisions about what constitutes value in the national accounts are made by blending different elements: anything that can be priced and exchanged legally; politically pragmatic decisions, such as accommodating technological change in the

computer industry or the embarrassingly large size of the financial sector; and the practical necessity of keeping the accounting manageable in very big and complex modern economies. This is all very well, but the fact that the production boundary debate is no longer explicit, nor linked openly to ideas about value, means that economic actors can – through sustained lobbying – quietly place themselves within the boundary. Their value-extracting activities are then counted in GDP – and very few notice.

Chapters 4, 5 and 6 examine the phenomenon of financialization: the growth of the financial sector and the spread of financial practices and attitudes into the real economy. In Chapter 4 I look at the emergence of finance as a major economic sector and its transition from being considered largely unproductive to becoming accepted as largely productive. As late as the 1960s, national accountants viewed financial activity not as generating value but as simply transferring existing value, which placed it outside the production boundary. Today, this view has changed fundamentally. In its current incarnation, finance is seen as earning profits from services reclassified as productive. I look at how and why this extraordinary redefinition took place, and ask if financial intermediation really has undergone a transformation into an inherently productive activity.

In Chapter 5 I explore the development of 'asset manager capitalism': how the financial sector expanded beyond the banks to incorporate an increasingly large number of intermediaries dedicated to managing funds (the asset management industry), and ask whether the role of these intermediaries, and the actual risks they take on, justify the rewards they earn. In doing so, I question the extent to which fund management and private equity have actually contributed to the productive economy. I ask, too, whether financial reform can be tackled today without a serious debate over whether activities in the financial sector are properly classified – are they what should be seen as rents, rather than profits? – and how we can go about making this distinction. If our national accounting systems are really rewarding value extraction as though it is value creation, maybe this can help us understand the dynamics of value destruction that characterized the financial crisis.

Building on this acceptance of finance as a productive activity, Chapter 6 examines the financialization of the whole economy. In seeking a quick return, short-term finance has affected industry: companies are run in the name of maximizing shareholder value (MSV). MSV arose in the 1970s as an attempt to revitalize corporate performance by invoking what was claimed to be the main purpose of the company: creating value for

shareholders. I will argue, however, that MSV has been detrimental to sustained economic growth, not least because it encourages short-term gain for shareholders at the expense of long-term gains for the company – a development closely linked to the increasing influence of fund managers seeking returns for their clients and for themselves. Underlying MSV is the notion of shareholders as the biggest risk takers, meriting the large rewards they often obtain.

Risk-taking is often the justification for the rewards investors reap, and Chapter 7 continues to look at other types of value extraction carried out in its name. Here I consider the kind of risk-taking required for radical technological innovation to occur. Innovation is without doubt one of the most risky and uncertain activities in capitalism: most attempts fail. But who takes it on? And what sort of incentives must be created? I explore the biased view of the current innovation narrative: how public-sector risk-taking is ignored, the state being seen as merely facilitating and 'de-risking' the private sector. The result has been policies, including reforms to the intellectual property rights (IPR) system, which have strengthened the power of incumbents, limiting innovation and creating 'unproductive entrepreneurship'.[20] Building on my previous book *The Entrepreneurial State*, I will show how entrepreneurs and venture capitalists have been hyped up to represent the most dynamic part of modern capitalism – innovation – and have presented themselves as 'wealth creators'. I will unpick the wealth-creating narrative to show how, ultimately, it is false. Claiming value in innovation, most recently with the concept of 'platforms' and the related notion of the sharing economy, is less about genuine innovation and more to do with facilitating value extraction through the capture of rents.

Picking up on the false innovation narrative, Chapter 8 will ask why the public sector is always described as slow, boring, bureaucratic and unproductive. Where did this depiction come from and who is benefiting from it? I will argue that, in the same way and at the same time that finance was made productive, the public sector has been made to appear unproductive. Modern economic thought has relegated government to just fixing market failures rather than actively creating and shaping markets. The value-creating role of the public sector, I contend, has been underestimated. The dominant view, which originated in the backlash against government in the 1980s, fundamentally affects how government sees itself: hesitant, cautious, careful not to overstep in case it should be accused of crowding out innovation, or accused of favouritism, 'picking winners'. In questioning why public-sector activities are ignored in GDP

accounting, I ask why this should matter, and outline what a different view of public value might look like.

It is, I conclude in Chapter 9, only through an open debate about value – its sources and the conditions that foster it – that we can help steer our economies in a direction that will produce more genuine innovation and less inequality, and which will also transform the financial sector into one that is truly focused on nurturing value creation in the real economy. It is not enough to critique speculation and short-term value extraction, and to argue for a more progressive tax system that targets wealth. We must ground those critiques in a different conversation about value creation, otherwise programmes for reform will continue to have little effect and will be easily lobbied against by the so-called 'wealth creators'.

This book does not try to argue for one correct theory of value. Rather, it aims to bring back value theory as a hotly debated area, relevant to the turbulent economic times in which we find ourselves. Value is not a given thing, unmistakably either inside or outside the production boundary; it is shaped and created. In my view, today finance nurtures not the industries for which it is meant to 'grease' the wheels of commerce, but rather other parts of the financial sector itself. It thus lies outside the boundary, even though it is formally counted as being inside. But this does not have to be the case: we can shape financial markets so that they do indeed belong inside the boundary. This would include both new financial institutions dedicated to lending to those organizations interested in long-term high-risk investments that can help foster a more innovative economy, as well as changing measures in the tax code that reward long-term investments over short-term ones. Similarly, as I discuss in Chapter 7, changes to the current unhelpful use of patents could help them stimulate innovation rather than stifle it.

To create a fairer economy, one where prosperity is more broadly shared and is therefore more sustainable, we need to reinvigorate a serious discussion about the nature and origin of value. We must reconsider the stories we are telling about who the value creators are, and what that says to us about how we define activities as economically productive and unproductive. We cannot limit progressive politics to taxing wealth, but require a new understanding of and debate about wealth creation so that it is more fiercely and openly contested. Words matter: we need a new vocabulary for policymaking. Policy is not just about 'intervening'. It is about shaping a different future: co-creating markets and value, not just 'fixing' markets or redistributing value. It's about taking risks, not only

‘de-risking’. And it must not be about levelling the playing field but about tilting it towards the kind of economy we want.

This idea that we can shape markets has important consequences. We can create a better economy by understanding that markets are outcomes of decisions that are made – in business, in public organizations and in civil society. The eight-hour working day has formed markets – and that was the result of a fight held in labour organizations. And perhaps the reason there is so much despair across the globe – despair now leading to populist politics – is that the economy is presented to us simply as ‘made’ by trade rules, technocrats and neoliberal forces. Indeed, as the book will show, ‘value’ theory itself is presented as a sort of objective force determined by supply and demand, rather than deeply embedded in particular ways of seeing the world. The economy can indeed be made and shaped – but it can be done either in fear or in hope.

The specific challenge I pose here is to move beyond Oscar Wilde’s cynic, who knows the price of everything but the value of nothing, towards an economics of hope, where we are better empowered to question the assumptions of economic theory and how they are presented to us. And to choose a different path among the many that are available.

1

A Brief History of Value

> *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.*
>
> Adam Smith, *The Wealth of Nations* (1776)

Today we take increasing prosperity for granted. We assume that by and large the next generation will be better off than the last. But it was not always so. For most of human history people had no such expectations and, partly because living standards improved at best very slowly, few thinkers devoted much time to asking why some economies grow and others do not. In the early modern period, the pace of change quickened. Previously static economies became dynamic. Movement was in the air. The rise of the nation state in Europe, the need to finance war, colonization, machinery, factories and coal, combined with expanding populations to stimulate new thinking across many fields. Governments and people of all stations in life wanted to know what was causing unprecedented movement and how it could be managed. What taxes can we raise? Why are my wages so low compared with the profits of capitalists? How sure can one be of the future when investing now? What creates value?

Understanding the nature of production is key to answering such questions. Once productive activities have been identified, economic policy can try to steer an economy, devoting a greater share of capital and effort to productive activities which propel and sustain economic growth. But the distinction between what is or is not productive has varied depending on economic, social and political forces. Ever since economists

began to explore the changing conditions of production some 300 years ago they have struggled to provide a rationale for labelling some activities productive and others unproductive. After all, economists are creatures of their time like everyone else; in terms of understanding value, what's important is to distinguish durable principles from transitory ones – and also, as we will see, the way that ideological positions develop.

This chapter explores how theories of value evolved from roughly the mid-seventeenth century to the mid-nineteenth century. The thinkers of the seventeenth century focused on how to calculate growth according to the needs of the time: fighting wars, or increasing competitiveness relative to another country – for example, England against its commercial and naval rival, Holland. The *mercantilists* focus on trade and the needs of merchants (selling things). From the mid-eighteenth to the late nineteenth century, economists saw value as arising from the amount of labour that went into production, at first farm labour (the *physiocrats*) and then industrial labour (the *classicals*). This value, they believed, therefore determined the price of what was finally sold. Their theories of value – of how wealth was created – were dynamic, reflecting a world being transformed socially and politically as well as economically. These economists focused on objective forces: the effects of changes in technology and the division of labour on how production and distribution are organized. Later, as we will see in the next chapter, they were superseded by another perspective – that of the *neoclassicals* – focused less on objective forces of production and more on the subjective nature of the 'preferences' of different actors in the economy.

THE MERCANTILISTS: TRADE AND TREASURE

Since ancient times, humanity has divided its economic activity into two types: productive and unproductive, virtuous and vile, industrious and lazy. The touchstone was generally what kind of activity was thought to further the common good. In the fourth century BC, Aristotle distinguished a variety of more or less virtuous jobs, depending on the class (citizen or slave) of the ancient Greek *polis* dweller.[1] In the New Testament, the apostle Matthew reported that Jesus said it was 'easier for a camel to go through the eye of a needle than for a rich man to enter into the Kingdom of God'.[2] During the Middle Ages, the Church disparaged and even denounced moneylenders and merchants who 'bought cheap and sold dear';[3] while they may not have been lazy, they were considered unproductive and vile.

Pre-modern definitions of what work was or was not useful were never clear-cut. With the onset of colonialism in the sixteenth century these definitions became even more blurred. European colonial conquest and the protection of trade routes with newly annexed lands were expensive. Governments had to find the money for armies, bureaucracies and the purchase of exotic merchandise. But help seemed to be at hand: extraordinary amounts of gold and silver were discovered in the Americas, and a vast treasure poured into Europe. As these precious metals represented wealth and prosperity, it seemed that whoever bought, owned and controlled the supply of them and the currencies minted from them was engaged in productive activities.

Scholars and politicians of the time who argued that accumulating precious metals was the route to national power and prosperity are called mercantilists (from *mercator*, the Latin word for merchant), because they espoused protectionist trade policies and positive trade balances to stimulate the inflow, and prevent the outflow, of gold and silver. The best-known English advocate of mercantilism was a merchant and director of the East India Company called Sir Thomas Mun (1571–1641). In his influential book *England's Treasure by Forraign Trade*, Mun summed up the mercantilist doctrine: we must, he said, 'sell more to strangers yearly than wee consume of theirs in value'.[4]

Mercantilists also defended the growth of national government as necessary to fund wars and expeditions to keep trade routes open and to control colonial markets. In England, Holland and France, mercantilists advocated shipping Acts, such as England's Navigation Act of 1651, which forced their countries' and colonies' trade exclusively into ships flying the national flag.

As mercantilist thinking developed, and people started to conceive of wealth production in national terms, the first estimates of national income – the total amount everyone in the country earned – started to appear. Seventeenth-century Britain saw two groundbreaking attempts to quantify national income. One was by Sir William Petty (1623–87), an adventurer, anatomist, physician and Member of Parliament, who was a tax administrator in Ireland under Oliver Cromwell's Commonwealth government.[5] The other was by the herald Gregory King (1648–1712), a genealogist, engraver and statistician whose work on enacting a new tax on marriages, births and burials provoked his interest in national accounting.

Petty and King were ingenious in their use of incomplete and messy data to generate surprisingly detailed income estimates. They had to work with rudimentary government tax figures, estimates of population and

patchy statistics on the consumption of basic commodities such as corn, wheat and beer. What their estimates lacked, however, was a clear value theory: Petty and King were concerned only with calculating the nation's output, not with how that output came about. Nevertheless, their attempts at national accounting were unprecedented and laid the foundations for modern national accounts.

In the 1660s, as Petty worked on his income studies, England was emerging from its experiment with republicanism, and was struggling with Holland and France for supremacy at sea. Petty wanted to find out whether England had the resources to survive these threats to its security: as he put it, to 'prove mathematically that the [English] State could raise a much larger revenue from taxes to finance its peace and wartime needs',[6] because he believed the country was richer than commonly thought.

Petty made a decisive breakthrough. He realized that income and expenditure at the national level should be the same. He understood that, if you treat a country as a closed system, each pound one person spends in it is another person's income of one pound. It was the first time anyone had grasped and worked with this fundamental insight. To make up for the lack of available statistics, Petty worked on the assumption that a nation's income is equal to its expenditure (omitting savings in good times, although he was aware of the potential discrepancy).[7] That meant he could use expenditure per person, multiplied by population, to arrive at the nation's income. In so doing he started, implicitly, to impose a production boundary, including within it only money spent on the production of 'Food, Housing, Cloaths, and all other necessaries'.[8] All other 'unnecessary expenses', as defined by Petty, were omitted.

In this way, by extension, Petty came to see any branch of the economy that did not produce those necessities as unproductive, adding nothing to national income. As he worked, his idea of the production boundary began to crystallize further, with 'Husbandmen, Seamen, Soldiers, Artizans and Merchants … the very Pillars of any Common-Wealth' on one side; and 'all the other great Professions' which 'do rise out of the infirmities and miscarriages of these' on the other.[9] By 'great professions' Petty meant lawyers, clergymen, civil servants, lords and the like. In other words, for Petty some 'great professions' were merely a necessary evil – needed simply for facilitating production and for maintaining the status quo – but not really essential to production or exchange. Although Petty did not believe that policy should be focused on controlling imports and exports, the mercantilists influenced him heavily. 'Merchandise', he argued, was more productive than manufacture and husbandry; the Dutch, he noted

approvingly, outsourced their husbandry to Poland and Denmark, enabling them to focus on more productive 'Trades and curious Arts'.[10] England, he concluded, would also benefit if more husbandmen became merchants.[11]

In the late 1690s, after the first publication of Petty's work *Political Arithmetick*, Gregory King made more detailed estimates of England's income. Like Petty, King was concerned with England's war-making potential and compared the country's income with those of France and Holland. Drawing on a wide variety of sources, he meticulously calculated the income and expenditure of some twenty different occupation groups in the country, from the aristocracy to lawyers, merchants to paupers. He even made forecasts, for example of population, predating the arrival of the forecasting 'science' some 250 years later, and estimated the crop yield of important agricultural items.

As in Petty's work, an implicit production boundary began to emerge when King assessed productivity, which he defined as income being greater than expenditure. King thought merchant traders were the most productive group, their income being a quarter more than their expenditure, followed by the 'temporal and spiritual lords', then by a variety of prestigious professions. On the boundary were farmers, who earned almost no more than they spent. Firmly on the 'unproductive' side were seamen, labourers, servants, cottagers, paupers and 'common soldiers'.[12] In King's view, the unproductive masses, representing slightly more than half the total population, were leeches on the public wealth because they consumed more than they produced.

Figure 2 shows that there were discrepancies between the 'productive' professions Petty and King identified. Almost all the professions Petty deemed unproductive King later saw as productive, while several of those producing value for Petty – seamen, soldiers and unskilled labourers – did not make the cut in King's analysis. Their different views may have stemmed from their backgrounds. A man of humble origins and republican instincts, Petty started out serving Oliver Cromwell; moving in aristocratic and court circles, King was perhaps less inclined to think that Petty's 'great professions' were unproductive. Both, however, classed 'vagrants' as unproductive, an analysis that has parallels today with people receiving welfare from governments financed by taxes on the productive sectors.

Some of Petty's and King's ideas have proved remarkably durable.[13] Perhaps most importantly, in what they both called 'Political Arithmetick' they laid the basis for what we today call the 'national accounts' to calculate GDP, the compass by which countries attempt to steer their national economic ships.

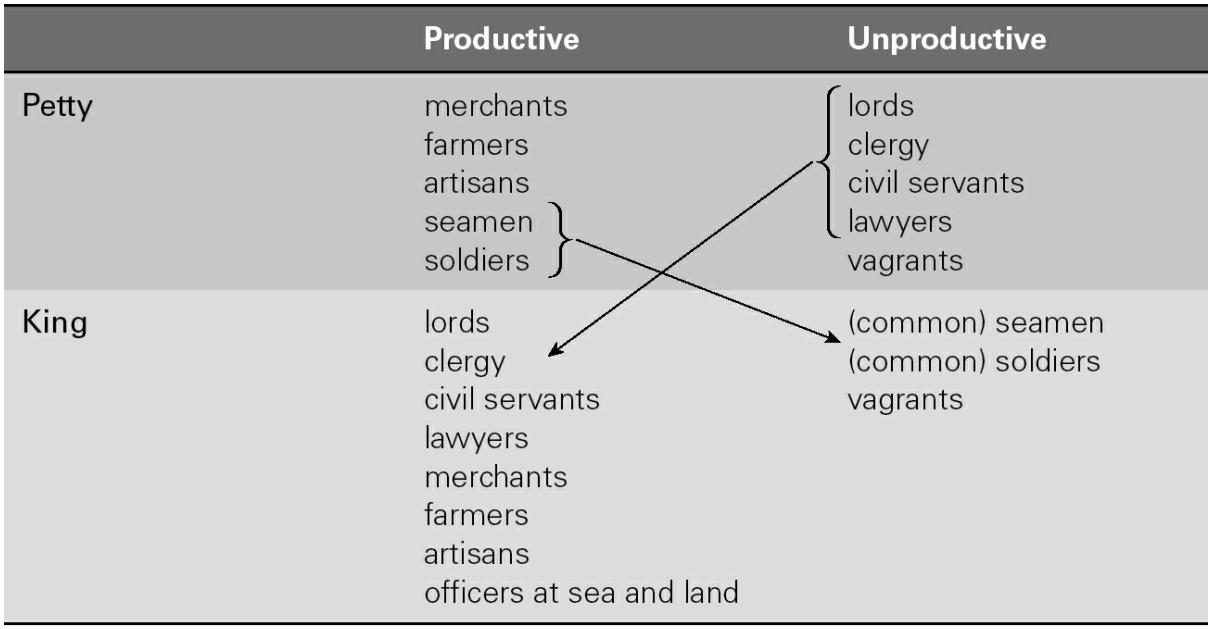


Figure 2. The production boundary in the 1600s

Mercantilist ideas still resonate in current economic practices. Modern ‘management’ of exchange rates by governments, trying to steal a competitive advantage for exports and accumulate foreign exchange reserves, harks back to mercantilist notions of boosting exports to accumulate gold and silver. Tariffs, import quotas and other measures to control trade and support domestic enterprises are also reminiscent of these early ideas about how value is created. There is basically nothing new in the calls to protect Western steel producers from Chinese imports or to subsidize domestic low-carbon energy generation to substitute for imports of oil, gas and coal. The emphasis by populist politicians on the negative effect of free trade, and the need to put up different types of walls to prevent the free movement of goods and labour, also gestures back to the mercantilist era, with emphasis more on getting the prices right (including exchange rates and wages) than on making the investments needed to create long-run growth and higher per capita income.

Petty and King were seminal figures in these early forays into the question of how and where value is created. Yet, ultimately, both could label productive and unproductive occupations however they chose. Their work was purely descriptive. It did not attempt to quantify or model relations between different groups and individuals in the economy,[14] or to quantify how the system reproduced itself and maintained the conditions for future production. In short, their work was not linked to an underlying theory of what constitutes wealth and where it comes from: a value theory. Any policy for economic growth was therefore idiosyncratic because it

was unclear what generated it. But during the following century, this would start to change.

As the study of economics developed during the course of the eighteenth century, thinkers became increasingly concerned with finding a theory to explain why some nations grew and prospered while others declined. Although the economists of the time did not use the term ‘production boundary’, the idea was at the heart of their work. Their search for the source of value led them to locate it in production, first in land – understandably so, in predominantly agrarian societies – and then, as economies became more industrialized, in labour. The labour theory of value reached its apogee with Karl Marx in the mid-nineteenth century, when the Industrial Revolution was in full swing.

THE PHYSIOCRATS: THE ANSWER LIES IN THE SOIL

The first efforts to find a formal theory of value came in the mid-eighteenth century from the court of Louis XV of France, in the twilight – so it turned out – of that country’s absolute monarchy. There, François Quesnay (1694–1774), often described as the ‘father of economics’, was the king’s physician and adviser. He used his medical training to understand the economy as a ‘metabolic’ system. Crucially, in metabolism, everything must come from somewhere and go somewhere – and that, for Quesnay, included wealth. Quesnay’s approach led him to formulate the first systematic theory of value that classified who is and is not productive in an economy, and to model how the entire economy could reproduce itself from the value generated by a small group of its members. In his seminal work *Tableau Économique*, published in 1758, he constructed an ‘economic table’ which showed how new value was created and circulated in the economy. In it he continued the metabolic analogy: pumps were drawn to signify the ways in which new value was introduced, and outgoing tubes illustrated how value left the system.

At the time Quesnay wrote, French society was already facing the problems that would lead to the French Revolution fifteen years after his death. French agriculture was in a bad state. Farmers were choked by high taxes, imposed by their usually noble landlords to fund their lavish lifestyles and by central government to finance war and trade. Adding to this burden, the French government’s mercantilist policy, faced with a now aggressively expanding Britain, kept the prices of agricultural produce low to provide cheap subsistence to domestic manufactures, which could in turn be cheaply made and exported in exchange for the highly coveted gold, still generally believed to be a measure of national wealth. Faced

with this situation, Quesnay and his followers built a powerful argument in favour of the farmers and against the mercantilists. Though they came to be known as the physiocrats, after one of Quesnay's publications, they called themselves something else: 'Les Économistes'.

Contrasting sharply with the prevailing mercantilist thinking that gave gold a privileged place, Quesnay believed that land was the source of all value. Figure 3 illustrates how for him, in the end, everything that nourished humans came from the earth. He pointed out that, unlike humans, Nature actually produced new things: grain out of small seeds for food, trees out of saplings and mineral ores from the earth from which houses and ships and machinery were built. By contrast, humans could not produce value. They could only transform it: bread from seeds, timber from wood, steel from iron. Since agriculture, husbandry, fishing, hunting and mining (all in the darker blob in Figure 3) bring Nature's bounty to society, Quesnay called them the 'productive class'. By contrast, he thought that nearly all other sectors of the economy – households, government, services and even industry, lumped together in the lighter blob – were unproductive.

Quesnay's classification was revolutionary. Breaking away from the mercantilists, who placed exchange and what was gained from it – gold – at the centre of value creation, he now linked value creation inextricably with production. Developing his classification of productive and unproductive work, Quesnay grouped society into three classes. First came farmers and related occupations working on the land and water; according to Quesnay, this was the only productive class. Next were manufacturers, artisans and related workers who transform the materials they receive from the productive class: wood and stone for furniture and houses, sheep's wool for clothing and metals from the mines for tools.[15] Yet, argued Quesnay, this class did not add value; rather, their work merely recirculated existing value. The third class was the unproductive 'proprietor', 'distributive' or 'sterile' class, which was made up of landlords, nobility and clergy. Here, 'distributive' was meant pejoratively: this class redistributes value, but only to itself, for the sole reason that it owns the land and does not give anything in return.[16]

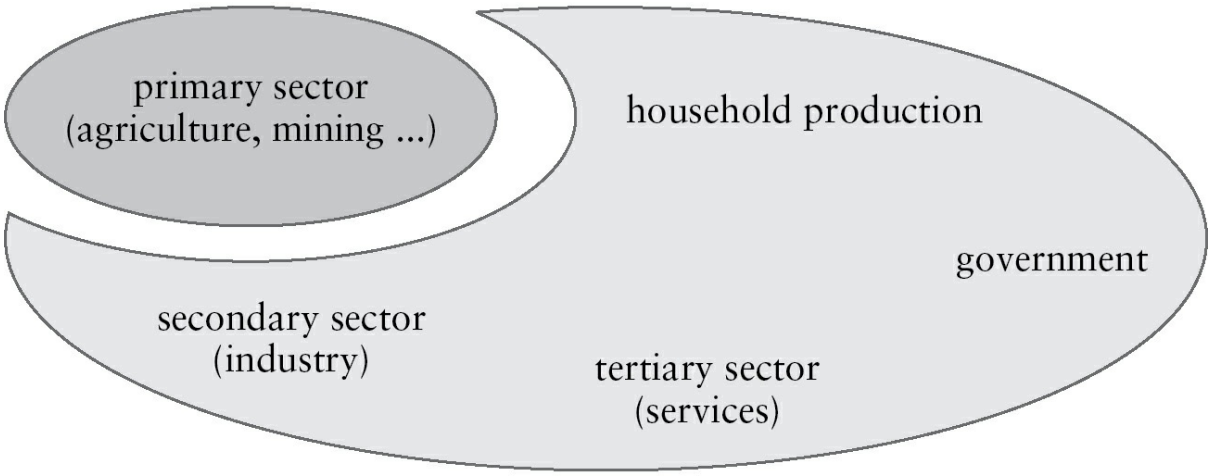
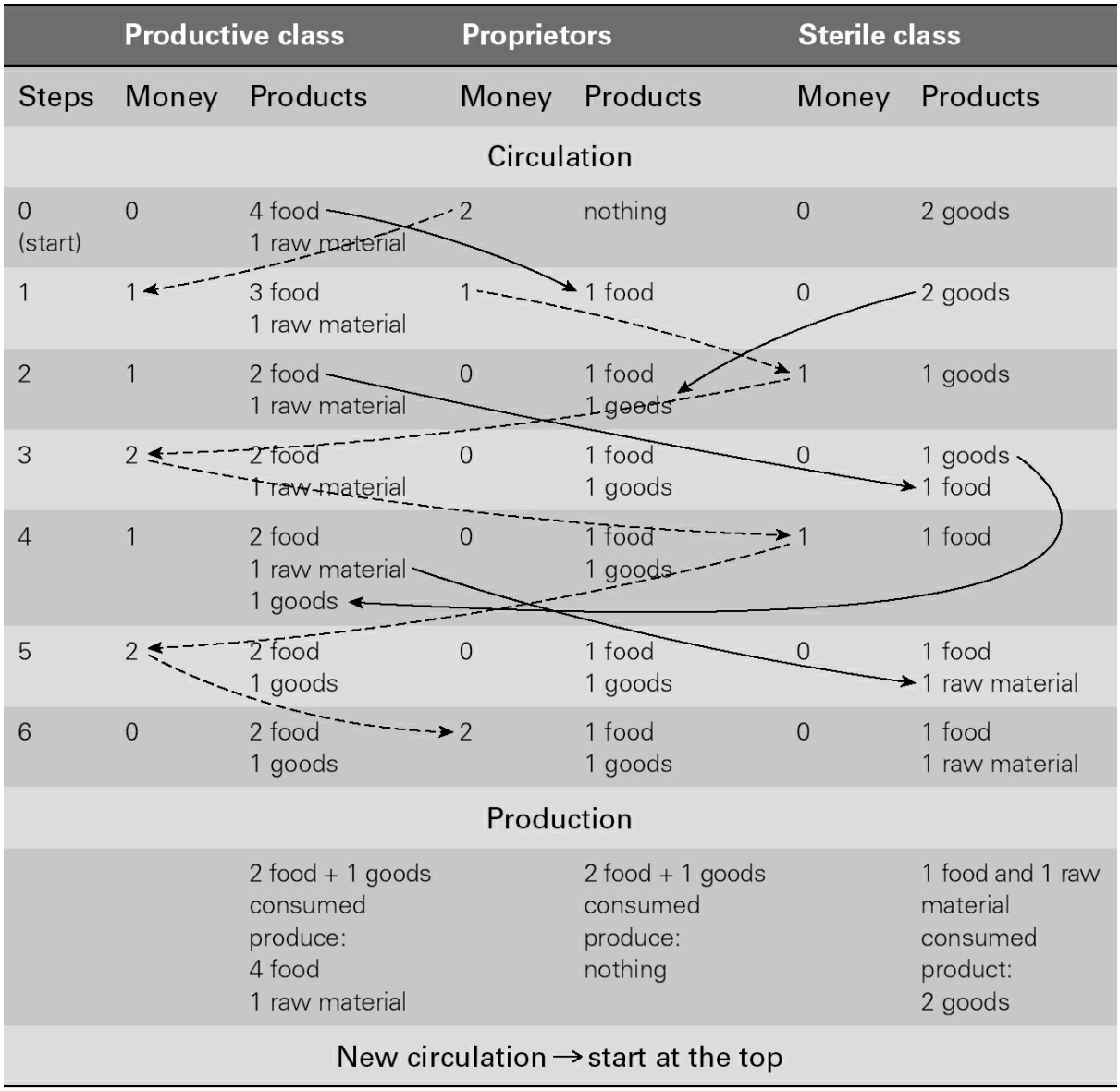


Figure 3. The production boundary in the 1700s

In Quesnay's table, the productive part of the system is entirely based on the farmers, but others also have a useful role in ensuring that the system reproduces itself. Figure 4 shows in detail the process of production, income and consumption of each class or economic sector, and how they interact. Perhaps the world's first spreadsheet, it is also the first consistent abstract model of economic growth.



	Productive class		Proprietors		Sterile class	
Steps	Money	Products	Money	Products	Money	Products
			Circulation			
0 (start)	0	4 food 1 raw material	2	nothing	0	2 goods
1	1	3 food 1 raw material	1	1 food	0	2 goods
2	1	2 food 1 raw material	0	1 food 1 goods	1	1 goods
3	2	2 food 1 raw material	0	1 food 1 goods	0	1 goods 1 food
4	1	2 food 1 raw material 1 goods	0	1 food 1 goods	1	1 food
5	2	2 food 1 goods	0	1 food 1 goods	0	1 food 1 raw material
6	0	2 food 1 goods	2	1 food 1 goods	0	1 food 1 raw material
			Production			
		2 food + 1 goods consumed produce: 4 food 1 raw material		2 food + 1 goods consumed produce: nothing		1 food and 1 raw material consumed product: 2 goods
			New circulation → start at the top			

Figure 4. Example of the *Tableau Économique*

A Numerical Example for the *Tableau Économique*

The logic of Quesnay's model is illustrated in Figure 4. The most important thing is where the initial wealth comes from, how it is circulated, and what percentage is reinvested into production (in nature) in the next round, creating more value – the latter being the essence of the growth process. In the simplest case of a non-expanding economy, the productive class has an initial amount of 'products of the earth' (translated from 'produits de la terre'), valued here for the sake of argument as 5 billion livres' worth. These are divided 4/5 food (for the farmers to subsist on) and 1/5 in material for the sterile class. The proprietors hold 2 billion in cash that they have

collected in taxes from the productive class, and the sterile class has an inventory of 2 billion livres' worth of tools and other manufactured goods.

From this, a process of circulation takes place, each step of which corresponds to a move from one row to the next. In every step, an equal amount of value changes hands, to prepare for the next round of production. But no new value is created. An exception is the step from period 5 to 6 in the circulation process, at which a transfer rather than exchange of 2 billion livres takes place. Only money flows, not products.[17] At the end, production takes place, with 2 billion surplus products in the productive sector, while 2 billion have been unproductively consumed in the proprietor class, starting a new round of circulation. Obviously, if the surplus is bigger than consumption, the economy will grow from round to round.

(All units are in billions of French livres; solid arrows indicate product flows, dashed arrows indicate money flows.)[18]

Most significant is how the table neatly shows, from row to row, that as long as what is produced is greater than what is consumed, an amount will be left over at the end to be reinvested, thereby allowing the economy to continue reproducing itself. If any of the unproductive members of society take too much, reducing the amount the farmer can reinvest in production, the economy will grind to a halt. In other words, if value extraction by the unproductive members exceeds value creation by the productive members, growth stops.

Though he himself did not use the term, Quesnay's theory of value incorporates a very clear production boundary, the first to be drawn with such precision, which makes it clear that the surplus the 'productive' sectors generate enables everyone else to live.

Other economists quickly weighed in with analysis and criticism of Quesnay's classification. Their attack centred on Quesnay's labelling of artisans and workers as 'sterile': a term that served Quesnay's political ends of defending the existing agrarian social order, but contradicted the everyday experience of a great number of people. Refining Quesnay's thinking, his contemporary A. R. J. Turgot retained the notion that all value came from the land, but noted the important role of artisans in keeping society afloat. He also recognized that there were other 'general needs' that some people had to fulfil – such as judges to administer justice – and that these functions were essential for value creation. Accordingly,

he re-labelled Quesnay's 'sterile' class as the 'stipendiary', or waged, class. And, since rich landowners could decide whether to carry out work themselves or hire others to do so using revenues from the land, Turgot labelled them the 'disposable class'. He also added the refinement that some farmers or artisans would employ others and make a profit. As farmers move from tilling the land to employing others, he argued, they remain productive and receive profits on their enterprise. It is only when they give up on overseeing farming altogether and simply live on their rent that they become 'disposable' rent collectors. Turgot's more refined analysis therefore placed emphasis on the character of the work being done, rather than the category of work itself.

Turgot's refinements were highly significant. In them, we see the emergent categories of wages, profits and rents: an explicit reference to the distribution of wealth and income that would become one of the cornerstones of economic thought in the centuries to come, and which is still used in national income accounting today. Yet, for Turgot, land remained the source of value: those who did not work it could not be included in the production boundary.[19]

Quesnay and Turgot's almost complete identification of productivity with the agricultural sector had an overriding aim. Their restrictive production boundary gave the landed aristocracy ammunition to use against mercantilism, which favoured the merchant class, and fitted an agricultural society better than an industrial one. Given the physiocrats' disregard for industry, it is hardly surprising that the most significant critique of their ideas came from the nation where it was already clear that value was not just produced in agriculture, but in other emerging sectors: a rapidly industrializing Britain. The most influential critic of all was Quesnay's contemporary, a man who had travelled in France and talked at length with him: Adam Smith.

CLASSICAL ECONOMICS: VALUE IN LABOUR

As industry developed rapidly through the eighteenth and nineteenth centuries, so too did the ideas of a succession of outstanding thinkers like Adam Smith (1723–90), David Ricardo (1772–1823) and Karl Marx (1818–83), a German who did much of his greatest work in England. Economists started to measure the market value of a product in terms of the amount of work, or labour, that had gone into its production. Accordingly, they paid close attention to how labour and working conditions were changing and to the adoption of new technologies and ways of organizing production.

In *The Wealth of Nations*, first published in 1776 and widely regarded as the founding work of economics, Smith's famous description of the division of labour in pin factories showed his understanding of how changes in the organization of work could affect productivity and therefore economic growth and wealth. Another enormously influential book, Ricardo's *On the Principles of Political Economy and Taxation*, first published in 1817, contained a famous chapter called 'On Machinery', in which he argued that mechanization was reducing demand for skilled labour and would depress wages. And in Marx's *Capital*, Volume 1 of which was first published in 1867, the chapter called 'The Working Day', which dealt with the development of the English Factory Acts governing working conditions, showed his fascination with production as the field on which the battle for workers' rights, higher wages and better conditions was being fought.

Smith, Ricardo and others of the time became known as the 'classical' economists. Marx, a late outrider, stands somewhat apart from this collective description. The word 'classical' was a conscious echo of the status given to writers and thinkers of the ancient Greek and Roman worlds, whose works were still the bedrock of education when the term 'classical economics' began to be used in the later nineteenth century. The classical economists redrew the production boundary in a way that made more sense for the period they lived in: one which saw the artisan-craft production of the guilds still prominent in Smith's time give way to the large-scale industry with huge numbers of urban workers – the proletariat – that Marx wrote about in the third quarter of the nineteenth century. Not for nothing was their emerging discipline called 'political economy'. It did not seem odd to contemporaries that economics was intimately part of studying society: they would have found odd the idea, widespread today, that economics is a neutral technical discipline which can be pursued in isolation of the prevailing social and political context. Although their theories differed in many respects, the classical economists shared two basic ideas: that value derived from the costs of production, principally labour; and that therefore activity subsequent to value created by labour, such as finance, did not in itself create value. Marx, we will see, was more subtle in his understanding of this distinction.

Adam Smith: The Birth of the Labour Theory of Value

Born in 1723 into a family of customs officials in Kirkcaldy, in the county of Fife, Scotland, Adam Smith became Professor of Moral Philosophy at

the University of Glasgow before turning his mind to what we now call economic questions, although at the time such questions were deeply influenced by philosophy and political thought.

With Britain well on the path to industrial capitalism, Smith's *The Wealth of Nations* highlighted the role of the division of labour in manufacturing. His account of pin-manufacturing continues to be cited today as one of the first examples of organizational and technological change at the centre of the economic growth process. Explaining the immense increase in productivity that occurred when one worker was no longer responsible for producing an entire pin, but only for a small part of it, Smith related how the division of labour allowed an increase in specialization and hence productivity:

> I have seen a small manufactory of this kind where ten men only were employed, and where some of them consequently performed two or three distinct operations. But though they were very poor, and therefore but indifferently accommodated with the necessary machinery, they could, when they exerted themselves, make among them about twelve pounds of pins in a day. There are in a pound upwards of four thousand pins of a middling size. Those ten persons, therefore, could make among them upwards of forty-eight thousand pins in a day. Each person, therefore, making a tenth part of forty-eight thousand pins, might be considered as making four thousand eight hundred pins in a day. But if they had all wrought separately and independently, and without any of them having been educated to this peculiar business, they certainly could not each of them have made twenty, perhaps not one pin in a day; that is, certainly, not the two hundred and fortieth, perhaps not the four thousand eight hundredth part of what they are at present capable of performing, in consequence of a proper division and combination of their different operations.[20]

These insights were original and profound. Smith was writing while the Industrial Revolution introduced machines into factories on a large scale. When harnessed to the division of labour, mechanization would radically increase productivity – the principal engine of economic growth. But even the simple reorganization of labour, without machinery, by which each worker specialized and developed skills in a specific area, enabled Smith to make this critical point.

Equally significant was Smith's analysis of how the 'market' determines the way in which consumers and producers interact. Such interaction, he contended, was not down to 'benevolence' or central planning.[21] Rather, it was due to the 'invisible hand' of the market:

> Every individual is continually exerting himself to find out the most advantageous employment for whatever capital he can command. It is his own advantage, indeed, and not that of the society which he has in view. But the study of his own advantage naturally, or rather necessarily, leads him to prefer that employment which is most advantageous to society … He intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was not part of his intention.[22]

Like Quesnay, Smith launched a more general attack on mercantilist policies which, he argued, restricted competition and trade. He also argued strongly for policies that would increase savings, and hence the amount of capital available for investment rather than unproductive consumption (say, on luxuries). But for Smith, industrial workers – not, as for Quesnay, farmers – were at the heart of the productive economy. Manufacturing labour, not land, was the source of value.[23] The labour theory of value was born.

Smith has become the figurehead of much modern economic theory because of his ideas about how capitalism is founded on supposedly immutable human behaviour, notably self-interest, and competition in a market economy. His metaphor of the 'invisible hand' has been cited *ad nauseam* to support the current orthodoxy that markets, left to themselves, may lead to a socially optimal outcome – indeed, more beneficial than if the state intervenes.

Smith's book is actually a collection of recipes for politicians and policymakers. Far from leaving everything to the market, he thinks of himself as giving guidance to 'statesmen' on how to act to 'enrich both the people and the sovereign'[24] – how to increase the wealth of nations. This is where Smith's value theory enters the picture. He was convinced that growth depended on increasing the relative share of 'manufactures' – factories employing formerly independent artisans or agricultural workers as dependent wage labourers – in the overall make-up of industry and believed that free trade was essential to bring this about. He felt that the enemies of growth were, first, the protectionist policies of mercantilists; second, the guilds protecting artisans' privileges; and third, a nobility that squandered its money on unproductive labour and lavish consumption. For Smith (as for Quesnay), employing an overly large portion of labour for unproductive purposes – such as the hoarding of cash, a practice that still afflicts our modern economies – prevents a nation from accumulating wealth.

Value, Smith believed, was proportional to the time spent by workers on production. For the purposes of his theory, Smith assumed a worker of average speed. Figure 5 shows how he drew a clear line (the production boundary) between productive and unproductive labour. For him, the boundary lay between material production – agriculture, manufacturing, mining in the figure's darker blob – and immaterial production in the lighter blob. The latter included all types of services (lawyers, carters, officials and so on) that were useful to manufactures, but were not actually involved in production itself. Smith said as much: labour, he suggested, is

productive when it is ‘realized’ in a permanent object.[25] His positioning of government on the ‘unproductive’ side of the boundary set the tone for much subsequent analysis and is a recurring theme in today’s debates about government’s role in the economy, epitomized by the Thatcher–Reagan reassertion in the 1980s of the primacy of markets in solving economic and social issues.

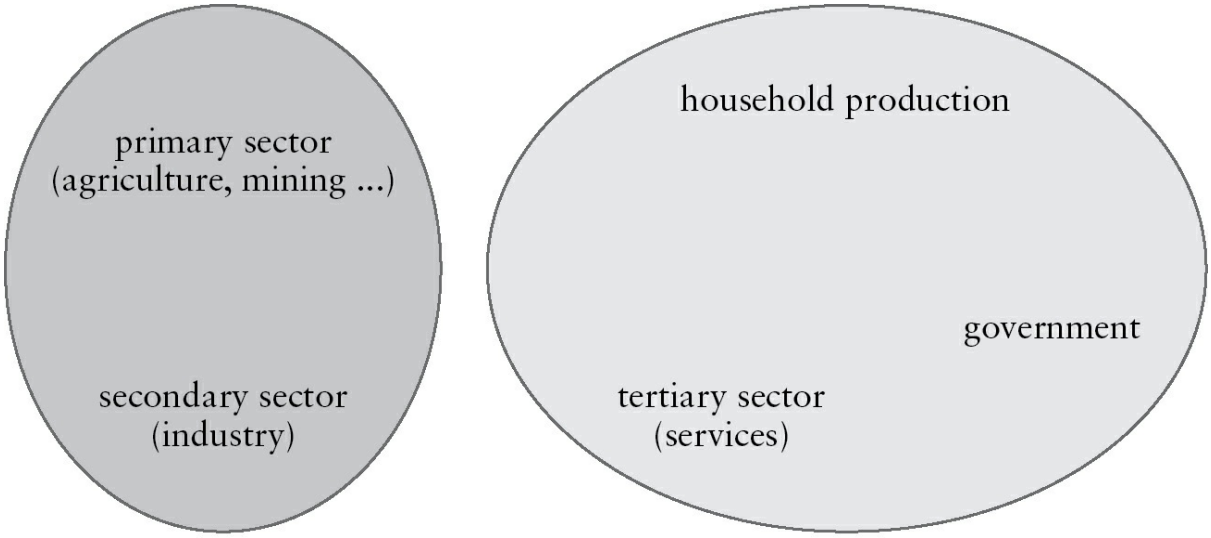


Figure 5. The production boundary according to Adam Smith

In Smith’s view, ‘how honourable, how useful, or how necessary soever’ a service may be, it simply does not reproduce the value used in maintaining (feeding, clothing, housing) unproductive labourers. Smith finds that even ‘the sovereign’, together with ‘all the officers both of justice and war who serve under him, the whole army and navy, are unproductive labourers’.[26] Priests, lawyers, doctors and performing artists are all lumped together as unproductive too.

What informs Smith’s classification is his conviction that some types of labour do not ‘reproduce’ the value needed to keep those workers alive at a subsistence level. In other words, if all the subsistence that was needed to keep a person alive was a certain amount of grain, then anyone who does not produce as much value as that amount of grain is by definition unproductive.

How, then, are those that do not produce this unit of value kept alive?

Smith’s answer lay in the concept of a ‘surplus’. Many productive workers produce the equivalent of more grain than they need to feed themselves to survive. A manufacturer makes things that, when exchanged, will yield more grain than needed to keep the productive workers alive. The surplus then sustains unproductive labourers, including

the entourages of aristocrats, who kept ‘a profuse and sumptuous table’ with ‘a great number of menial servants, and a multitude of dogs and horses’.

This is where Smith addressed head-on how the wealth of nations could grow. It was in effect his policy advice. Instead of ‘wasting’ the surplus on paying for unproductive labour, he argued, it should be saved and invested in more production so that the whole nation could become richer.[27] Smith was not criticizing the wealthy per se. But he was criticizing those who wasted their wealth on lavish consumption – ‘collecting books, statues, pictures’, or ‘more frivolous, jewels, baubles, ingenious trinkets’ – instead of productive investment. (This, after all, was the age of the Grand Tour, when young aristocrats travelled to the Continent to improve their education and returned laden with ancient artefacts.) Smith was particularly attracted to the prospect of investment in machines, then just beginning to be used in factories, because they improved workers’ productivity.

His emphasis on investment linked directly to his ideas about rent. Smith believed that there were three kinds of income: wages for labour in capitalist enterprises; profits for capitalists who owned the means of production; and rents from ownership of land. When these three sources of income are paid at their competitive level, together they determine what he called the ‘competitive price’.[28] Since land was necessary, rent from land was a ‘natural’ part of the economy. But that did not mean rent was productive: ‘the landlords, like all other men, love to reap where they never sowed and demand a rent [from the earth] even for its natural produce’.[29] Indeed, Smith asserted, the principle of rent from land could be extended to other monopolies, such as the right to import a particular commodity or the right to plead at the bar. Smith was well aware of the damage monopolies could do. In the seventeenth century, a government desperate for revenue had granted – often to well-placed courtiers – an extraordinary range of monopolies, from daily necessities such as beer and salt to mousetraps and spectacles. In 1621 there were said to be 700 monopolies, and by the late 1630s they were bringing in £100,000 a year to the Exchequer.[30] But this epidemic of rent-seeking was deeply unpopular and was choking the economy: more than that, it was one of the proximate causes of the Civil War, which led to the execution of Charles I. Many Englishmen understood what Smith meant when he said that a free market was one free of rent.

Smith’s penetrating analysis of how advanced capitalist economies functioned won him many followers. Equally, his staunch advocacy of free

trade, in an era in which mercantilist policies were beginning to be seen as old-fashioned (Smith, indeed, believed that merchants were unproductive because they only provided the ephemeral service of moving goods around, rather than producing anything of value), made his book a hit among the ‘free traders’ who eventually overturned England’s Corn Laws, which imposed heavy tariffs on imported corn to protect domestic landowners, and other protectionist measures. Armed with Smith’s ideas, free traders showed that nations could get richer even if there was no trade surplus and no gold accumulation. Amassing gold was unnecessary and insufficient for growth. Huge amounts of gold flowed to Spain from its colonies, but the kingdom did not become more productive.

The victory of the free traders over the mercantilists is better understood in terms of their rival conceptions of value. Mercantilists thought gold had inherent worth and that everything else could be valued in terms of how much gold it was exchanged for. Following Smith, free traders could trace value to labour, and the logic of value was thereby inverted. Gold, like all other things, was valued by how much labour it took to produce.[31]

Smith’s theory was not immune to criticism. He had actually put forward at least two theories of value, which created confusion about both the production boundary and precisely who was productive – in particular, whether the provision of services in themselves created value.[32]

In essence, Smith was confused about the distinction between material and immaterial production. For Smith, as we have seen, a servant ‘adds’ no value that could be used by the master on something other than, literally, keeping the servant alive. But he also argued that if a manufacturing worker earns £1 in turning a quantity of cotton, whose other inputs also cost £1, into a piece of cloth that sells for £3, then the worker will have repaid his service and the master has made a profit of £1. Here a definition of productivity, irrespective of whether what is produced is a solid product or a service, emerges. Adding value in any branch of production is productive; not adding value is unproductive. Following this definition, services such as cleaning or vehicle repair can be productive – thereby invalidating Smith’s own material–immaterial division of the production boundary. The debates about Smith’s theories of value rumbled on for centuries. Other of Smith’s ideas, such as free trade and the unproductive nature of government, have also left an enduring legacy.

But he is often misconstrued. His understanding of politics and philosophy was never sidelined in his economic reasoning. His *Theory of Moral Sentiments* and *The Wealth of Nations* were not contradictory but part of his deep analysis of what drives human behaviour and how

societies organize themselves, and why some societies might grow in wealth more than others. Smith's analysis of 'free markets' was closely tied to his understanding of production, and the need to limit rent-seeking behaviour.

David Ricardo: Grounding Smith's Value Theory

In the 1810s, another towering figure of the English classical economic school used the labour theory of value and productiveness to explain how society maintains the conditions which enable it to reproduce itself. David Ricardo came from a Sephardic Jewish family which originated in Portugal and moved to Holland before settling in England. Ricardo followed his father as a London stockbroker, although he was later estranged from his family after becoming a Unitarian. He grew fabulously rich from his speculative activities, most notoriously by profiting from inaccurate information that was circulating on the Battle of Waterloo in 1815. He was said to have made £1 million (in 1815 value) from holding on to bonds while everyone else was selling (due to the false rumours that Wellington was losing against Napoleon), an almost unimaginable sum at the time, after which he promptly and wisely retired to the country, well away from London.

Ricardo was drawn to economics by reading Smith's *The Wealth of Nations*, but was concerned with something that he felt was glaringly absent from Smith's theory of value: how that value was distributed throughout society – or what we would today call income distribution. It need hardly be said that, in today's world of growing inequality of income and wealth, this question remains profoundly relevant.

Smith had observed that the value produced by labour, when sold, is redistributed as wages, profits and rent; he had also seen that labour's exact share of this value – wages – would vary.[33] However, Smith had no coherent explanation for the way in which wages were apportioned, or why they differed between professions and countries or over time.[34] Ricardo, by contrast, felt that the distribution of wages was, as he stressed in his magnum opus *On the Principles of Political Economy and Taxation*, the 'principle problem' in economics and ultimately regulates the growth and wealth of a nation.

Ricardo actually believed in the labour theory of value, and, unlike Smith, was at pains to point out that the value of a commodity was strictly proportional to the amount of labour time needed to produce it. Ricardo emphasized agriculture for a different reason from Quesnay. He wanted to

explain the distribution of income, and for him productivity in agriculture was the hinge upon which that distribution turned. Workers, Ricardo believed, were paid a subsistence wage: in essence, they earned enough to pay for food and shelter. But food comes from agriculture, so the price of food regulates wages: a low price of food (or 'corn', as Ricardo wrote in the language of the day) will permit lower wages and therefore higher profits and incentives to invest in future production (for example in manufacturing) and promote economic growth. A high wage due to low productivity in agriculture will mean lower profits, and hence little investment in future production, which in turn leads to slower economic growth.

Ricardo inherited this 'dismal theory' of wages from his contemporary Thomas Malthus (1766–1834), another English writer on political economy, who proposed that whenever real wages are above subsistence level, the population will grow until it is so large that the demand for food will push up food prices enough to bring wages back to subsistence level.[35]

In Ricardo's view, then, wages depended heavily on the productivity of agriculture: if productivity rose and food became cheaper, wages would fall. And in manufacturing and the other branches of the economy, whatever did not have to be paid to the worker would flow to the capitalist as profit. Profits are the residual from the value that workers produce and do not need to consume for their own 'maintenance', as Ricardo put it, 'to subsist and perpetuate their race'.[36]

This in turn leads to Ricardo's theory of growth and accumulation – increasing the stock of capital or wealth to help fuel subsequent further increases in wealth. As profits grow, so capitalists invest and expand production, which in turn creates more jobs and raises wages, thereby increasing the population, whose wages finally go back to subsistence level, and so on. The economy is a perpetual growth machine, with more and more people earning the subsistence wage.

But Ricardo's theoretical genius really came to the fore in tackling his third class of society: landlords. Production in agriculture depends on two types of input: goods and services needed for production. One type can be scaled – increased in proportion to requirements. It includes labour, machinery, seeds and water. The other type cannot be scaled: good arable land. As Mark Twain is supposed to have said, 'Buy land, they're not making it any more.'

Since the population will grow thanks to investment and rising wages, and more and more food will need to be produced to feed everyone, at some point all the best land for corn production will be spoken for. Less

fertile or productive land will then be cultivated. However, since all the corn is sold at one price to the workers, who are on subsistence wages, the more productive land already in use yields a higher profit than the less productive land. Here Ricardo developed his celebrated theory of rent.

Ricardo defined rent as a transfer of profit to landlords simply because they had a monopoly of a scarce asset. There was no assumption, as in modern neoclassical theory (reviewed in Chapter 2), that these rents would be competed away. They remained due to power relationships inherent in the capitalist system. In Ricardo's time much of the arable land was owned by aristocrats and landed gentry but worked by tenant farmers or labourers. Ricardo proposed that the rent from more productive land always goes to the landlord because of competition between tenants. If the capitalist farmer – the tenant – wants to hang on to the largest possible profit by paying less rent, the landlord can give the lease to a competing farmer who will pay a higher rent and therefore be willing to work the land for only the standard profit. As this process goes on, land of increasingly poor quality will be brought into production, and a greater portion of the income will go to the landlords. Ricardo predicted that rents would rise.

More significantly, rising rents were the flipside of rising food prices, caused by lack of good-quality agricultural land. More costly food increased the wages workers needed for subsistence. This growing wage share, Ricardo believed, put a squeeze on profits in other sectors such as manufacturing. As economic development proceeded, the profit rate – basically the manufacturing capitalist's return on capital – would fall. The profit share – the part of the national income going to capitalists – would also fall. Correspondingly, the wage share going to manufacturing workers would rise. But the extra wages would have to be spent on food, which was more expensive because landlords were charging higher rents. As a result, much of the nation's income would ultimately go to landlords. This would halt further economic growth and investment in, say, manufacturing because the low returns would not justify the risks.[37]

By highlighting the different types of incomes earned, such as rent, profits and wages, Ricardo drew attention to an important question. When goods are sold, how are the proceeds of that sale divided? Does everyone involved get their 'just share' for the amount of effort they put into production? Ricardo's answer was an emphatic 'No'.

If some input into production – such as good arable land – is scarce, the cost of producing the same output – a given quantity of corn – will vary according to availability of the input. The cost is likely to be lower with good land, higher with inferior land. Profits, instead, are likely to be higher

with good land and lower with inferior land. The owner of good land will pocket the difference in profit between the good land and inferior land simply because he or she has a monopoly of that asset.[38] Ricardo's theory was so convincing that it is, in essence, still used today in economics to explain how rents work.[39] Rents in this sense could mean a patent on a drug, control of a rare mineral such as diamonds, or rents in the everyday sense of what you pay a landlord to live in a flat. In the modern world, oil producers like those of the Organization of Petroleum Exporting Countries (OPEC) collect rents from their control of an essential resource.

Ricardo's gloomy picture of economic stagnation is relevant to a modern debate: how the rise of the financial sector in recent decades and the massive rents it earned from speculative activity have created disincentives for industrial production. Some heterodox economists today argue that growth will fall if finance becomes too big relative to the rest of the economy (industry) because real profits come from the production of new goods and services rather than from simple transfers of money earned from those goods and services.[40] To 'rebalance' the economy, the argument runs, we must allow genuine profits from production to win over rents – which, as we can see here, is exactly the argument Ricardo made 200 years ago, and John Maynard Keynes was to make 100 years later.[41]

Indeed, as is also argued today, Ricardo believed that the pool of (mainly unskilled) workers held the losing ticket. In Ricardo's day, agricultural labour flocked to the fast-growing cities and the supply of unskilled labour exceeded demand for it. Without bargaining power, these workers were paid a meagre subsistence wage. Ricardo's portrayal of rents dominating production also had a political impact. It helped to persuade Britain to abolish the Corn Laws in 1846 and embrace free trade, which diminished the power of big vested interests and allowed production costs, rather than embedded monopoly and the privileges that went with it, to govern production. The ensuing decades saw Britain become the 'workshop of the world'. But the abolition of the Corn Laws brought about a political transformation as well as an economic one: it tipped the balance of power away from aristocratic landlords and towards manufacturing as the nineteenth century wore on. Value theory influenced political behaviour, and vice versa – the performativity referred to in the Preface.

Other lessons about the sources of value and who generates it can be drawn from Ricardo's model of accumulation. Like Smith, Ricardo was concerned with understanding how the economy reproduces itself. Like Smith, he focused on the difference between investment in durable capital and consumption: 'When the annual productions of a country more than

replace its annual consumption, it is said to increase its capital; when its annual consumption is not at least replaced by its annual production, it is said to diminish its capital.'[42] Ricardo hastened to add, though, that all goods produced – from clothes to carts – must be consumed or used; otherwise they would depreciate just like inventory.

Here Ricardo made a fundamental point about consumption, by which he means consumption by capitalists, not just households. As with production, consumption can be productive or unproductive. The productive kind might be a capitalist who 'consumes' his capital to buy labour, which in turn reproduces that capital and turns a profit. The alternative – unproductive consumption – is capital spent on luxuries that do not lead to reproduction of that capital expenditure. On this matter, Ricardo is absolutely clear: 'It makes the greatest difference imaginable whether they are consumed by those who reproduce, or by those who do not reproduce another value.'[43]

So Ricardo's heroes are the industrial capitalists, 'those who reproduce', who can ensure that workers subsist and generate a surplus that is free for the capitalist to use as he or she sees fit. His villains are those 'who do not reproduce' – the landed nobility, the owners of scarce land who charge very high rents and appropriate the surplus.[44] For Ricardo, capitalists would put that surplus to productive use, but landlords – including the nobility – would waste it on lavish lifestyles. Ricardo echoes Smith here. Both had seen with their own eyes the extravagance of the aristocracy, a class which often seemed better at spending money than making it and was addicted to that ultimate unproductive activity – gambling. But Ricardo parted company from Smith because he was not concerned about whether production activities were 'material' (making cloth) or 'immaterial' (selling cloth). To Ricardo, it was more important that, if a surplus was produced, it was consumed productively.

Significantly for our discussion, Ricardo singled out government as the ultimate example of unproductive consumption. Government, in his view, is a dangerous leech on the surplus. Most of government spending comes from taxes, and if it consumes – by spending on armies, for example – too large a share of the national income, 'the resources of the people and the state will fall away with increasing rapidity, and distress and ruin will follow'.[45] Ricardo believed that government is by nature unproductive.

At the time Ricardo was writing, such issues were uppermost. Only a few years earlier, the British government had had to raise unprecedented amounts of money from taxes and issuing bonds to wage the war against Napoleon, from which the nation emerged heavily in debt. Could it afford

the immense military expenditure which Ricardo's theory deemed unproductive? He found to his relief that the increase in value production by private companies more than compensated for the increase in unproductive government consumption. Unlike Smith, Ricardo did not write about that part of government expenditure which creates the conditions for productivity in the first place: infrastructure (bridges, roads, ports and so on), national defence and the rule of law. By omitting to discuss the role of government in productivity, he paved the way for generations of economists to be equally oblivious – with hugely significant consequences that we will look at in Chapter 8.

In essence, Ricardo's theory of value and growth led to a production boundary that does not depend on a job or profession itself (manufacturer, farmer or vicar) or on whether the activity is material or immaterial. He believed that industrial production in general leads to surpluses, but for him the real question is how those surpluses are used. If the surpluses finance productive consumption, they are productive; if not, they are unproductive.

Ricardo focused on the 'plight' of capitalists and their struggle against landlords. However, he never addressed the awkward fact that labour creates value but the capitalists get the spoils – the surplus over and above the subsistence wages paid to labourers. In the course of the nineteenth century, as England industrialized, inequalities and injustices multiplied. The labour theory of value was to interpret production in a way that cast capitalists in a much less favourable light.

Karl Marx on 'Production' Labour

Ricardo's appreciation of the dynamism of capitalism compared with past eras prefigures the emphasis Marx placed a generation later on the system's unprecedented power to transform societies. Born in 1818, Marx grew up in the German city of Trier, one of nine children of Jewish parents, both lawyers. In his own legal studies at university, Marx was drawn to a critical version of Hegel's philosophy of dialectics, propounded by Hegel's disciples, which set out how intellectual thought proceeds via negation and contradiction, through a thesis, its antithesis, and then a synthesis. Marx was particularly interested in how history is shaped by contradictions between material forces – such as capital and labour – and by the resolution or synthesis of those contradictions. After being barred from taking a professorship at the University of Jena because of his radical political leanings, he became editor of a progressive newspaper,

Rheinische Zeitung. Then in 1843 he moved to Paris, where he met Friedrich Engels, his future co-author and collaborator. Two years later Marx was expelled from France because of his socialist political activities and settled in Brussels. There in 1848 he published with Engels the *Communist Manifesto*. Marx wrote voluminously on politics for the rest of his life but it is remarkable that, despite being opposed to capitalism, he analysed it objectively in order to understand where it was taking humankind and what the alternatives might be.

Marx developed his own version of the labour theory of value. He emphasized how definitions of 'productive' activity depend on historical circumstances – the society of any given time. He also focused on the nature of productive activity within the capitalist system. Under capitalism, firms produce commodities – a general term for anything from nuts and bolts to complete machines. If commodities are exchanged – sold – they are said to have an exchange value. If you produce a commodity which you consume yourself it does not have an exchange value. Exchange value crystallizes the value inherent in commodities.

The source of that inherent value is the one special commodity workers own: their labour power, or – put another way – their capacity to work. Capitalists buy labour power with their capital. In exchange, they pay workers a wage. Workers' wages buy the commodities such as food and housing needed to restore a worker's strength to work. In this way, wages express the value of the goods that restore labour power.

This description of the source of value largely followed Ricardo. But Ricardo had tried unsuccessfully to find an external commodity that could serve as an 'invariable standard of value' by which the value of all other products could be determined. Marx solved this problem by locating this invariable measure in workers themselves. He was careful to distinguish labour expended in production from labour power, which is the *capacity* to work. Workers expend labour, not labour power. And in this distinction lies the secret of Marx's theory of value. Humans can create more value than they need to restore their labour power. For instance, if a worker has to work five hours to produce the value needed to restore labour power per day, the labour power's value is equivalent to the five hours of work. However, if the working day lasts ten hours, the additional five hours' work will create value over and above that needed to restore labour power. Labour power creates surplus value.

The ingenuity of capitalism, according to Marx, is that it can organize production to make workers generate unprecedented amounts of this surplus value. In early societies of hunter-gatherers and subsistence

farmers, people worked enough to create the value that would allow them to survive, but no surplus over and above that. Later, under feudalism, they could be forced to produce enough surplus to satisfy the (unproductive) consumption of the feudal lord, which, as Smith and Ricardo knew, could be substantial. But after the means of production were taken away from independent producers – mostly by violence and expropriation through property rights legislation, such as enclosures of common land in England by big landowners – they became workers, 'free' and without property.

Capitalists were able to purchase the workers' labour power because workers lost their independent means of subsistence and needed a wage to survive. The trick is to get them to work longer than needed to produce value (wages) that they spend on their subsistence needs – again, food and housing.[46] Workers, in other words, are *exploited* because capitalists pocket the surplus value workers produce over and above their subsistence requirements. And, unlike the feudal lords, capitalists will not squander all of the surplus on consumption, but will have incentives to reinvest part of it in expanding production to make yet more profits. However, Marx noted that there was a contradiction in the system. The drive to increase productivity would increase mechanization, which, in displacing labour (machines taking over human work), would then eventually reduce the key source of profits: labour power. He also foresaw the problem of growing financialization, which could potentially undermine industrial production. Throughout his analysis, his focus was on change, and the effects of change on the creation of value.

Indeed, the extraordinary aspect of Marx's theory is his fundamental insight that capitalism is dynamic and constantly changing. But it was not just economically dynamic. Marx was struck by the social upheavals he could see all around him, such as the mass movement of rural workers into cities, which created an urban proletariat. He saw that capitalist *society*, not just the capitalist economy, was utterly different from preceding societies and was in permanent flux – a very evident phenomenon today as we struggle to come to terms with the massive changes brought by digital, nano, biological and other technologies.

Economists had previously thought of 'capital' as purely physical – machinery and buildings, for example – and surplus as solely positive, helping the economy to reproduce itself and grow. But Marx gives capital a social dimension and surplus a negative connotation. Labour produces surplus value, which fuels capital accumulation and economic growth. But capital accumulation is not just due to productive labour. It is also deeply social. Because workers do not own the means of production they are

‘alienated’ from their work. The surplus they produce is taken away from them. Work is necessary for earning the wages they receive to buy the food, shelter and clothes they need to survive.[47] Moreover, in a capitalist market society, relations between people are mediated by commodity exchange. In a specialized society with division of labour, humans produce the social product – net national income – together and depend on other humans. But precisely because the division of labour, which Smith extolled, left most workers overly specializing in discrete aspects of the production process, he believed that social relations became relations between commodities (things).[48]

Marx was so fascinated by the dynamics of capitalism that he produced his own theory of value to explain how it works. Unlike earlier economists, who tended to define production by sector or occupation (agriculture or manufacturing, merchant or clergyman), Marx defined the production boundary in terms of *how* profits are made. Marx asked how, by owning the means of production, the capitalist could appropriate surplus value while the workers who provided the labour received barely enough to live on – exactly the question Big Bill Heywood posed. By placing this distinction at the heart of value theory, Marx generated a new and unprecedented production boundary. Marx’s value theory changed economics – at least for a time.

Marx argued that workers are productive if they create surplus value which the capitalist class then retains. For Marx, while workers in capitalist production are productive, the key questions when drawing his production boundary are: who participates in capitalist production? And who receives the surplus that is produced?

Figure 6 gives a graphical answer to these questions. The production sphere, the light grey blob, includes three basic sectors: primary, comprising essential materials such as food and minerals (the only source of value for Quesnay); secondary, which is industry, the basis of value creation in Smith and Ricardo; and tertiary, the services considered by Smith to be ‘immaterial’. The darker blob within, called the ‘circulation sphere’, reflects Marx’s analysis, which we will discuss later, that some aspects of finance are essential to production and deserve to be placed on that side of the production boundary. On the other side of the boundary, Marx followed Smith and Ricardo in regarding government and households as unproductive.

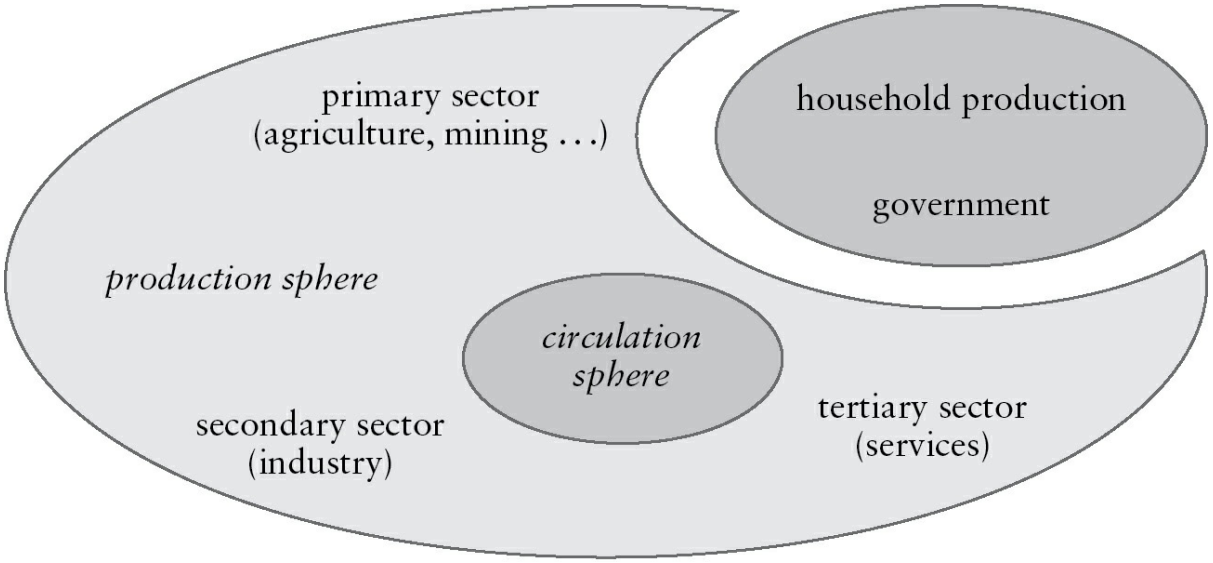


Figure 6. The production boundary according to Karl Marx

At any moment in a capitalist economy, there is a ratio of surplus value to value used for workers' subsistence – what Marx calls simply the rate of surplus value. It determines what share of the economic product can potentially be used for accumulation and growth. Marx referred to capital that is used to hire labour as 'variable' capital: the workers produce more capital than is invested in them, so the capital that hires them 'varies' in relation to the capitalist's total capital. Capital not used to hire workers is invested in other means of production that are 'constant' capital – including machinery, land, buildings and raw materials – whose value is preserved but not increased during production.[49]

The value used for workers' subsistence, the 'wage share', could not be less than was needed to restore labour power or workers would perish, leaving the capitalist unable to produce surplus value. Historically, the wages of the poor had tended to be at subsistence level. But here Marx introduces a powerful new idea which has informed thinking ever since: class struggle. Workers' wages were set by class struggle. The side with more power could force through a wage rate favourable to itself. Which class had more power was related to what we would call today the tightness of the labour market. If wages increased because workers had a lot of bargaining power in a tight labour market, capitalists would substitute more machines for labour, creating more unemployment and competition among workers for jobs. Marx thought that capitalists would try to keep a 'reserve army' of the unemployed to hold down wages and maintain or increase their own share of the value workers created.

The value of labour power is expressed to workers as wages, to capitalists as profits. The rate of profit for an enterprise is the surplus value divided by variable and constant capital – roughly what today we call the rate of return on a company's assets. The average profit rate of the economy as a whole is total surplus value divided by total variable and constant capital. But the size of the average profit rate depends on the composition of capital (how much variable and constant capital) and on class struggle – effectively, the size of workers' wages relative to value produced. The average profit rate is also affected by economies of scale as the productivity of workers rises with a growing market and the increasing specialization of workers.[50] In particular, Marx believed that increasing agricultural production would not lead to Ricardo's stationary, food-constrained world.[51] He was right: broadly speaking, food production has kept pace with population increase. Marx was also acute in his understanding of the capacity of technology to transform society. He would not have been surprised by the extent to which automation has replaced people, nor perhaps by the possibility of machines more intelligent than their human creators.

Marx's analysis of who got what in capitalism did not stop there. He also distinguished between different functions of various capitalist actors in the economy. In doing so he used his value theory shrewdly to identify those who produce value and those who do not.

Like economists before him, Marx believed that competition would tend to equalize rates of profits across the economy.[52] But at this point Marx introduced a distinction that is critically important for his and for subsequent theories of value: the way in which different kinds of capitalists came by their profits. The first two categories Marx identified were production (or industrial) capital and commercial capital. The first produces commodities; the second circulates commodities by selling them, making the money received available to production capital for buying the means of production (the dark grey sphere in the lighter blob in Figure 6). As Marx explained, the first creates surplus value, the second 'realizes' it. Any unsold commodity will therefore be of no use to a capitalist, regardless of how much he or she exploits his or her workers, because no surplus value is realized. Commercial capital, Marx noted, had existed for millennia: international merchants such as the Phoenicians and the Hanse bought cheap and sold dear. What they did not do was to add value by capitalist production. Under capitalism, the commercial capitalists realize the value produced by the production capitalists. To apply Marx's theory to a modern-world example, Amazon is a commercial capitalist because it

is a means by which production capitalists sell their goods and realize surplus value. Banks' money transfer services are also an example of commercial capital.[53]

Marx suggested that, initially, production enterprises might also carry out commercial capital activities. As production expands, however, separate capitalist enterprises will probably emerge to carry out these functions as commodity or money capitalists. Crucially, these capitalists and the labour they employ are purely concerned with the 'circulation' of capital; they do not produce commodities which generate surplus value and therefore they are unproductive.[54] However, because they are also capitalist firms, they require the same rate of profit as does production capital. Consequently, some surplus value is diverted to become their income, diminishing the average profit rate in the economy.[55] Although labour in firms engaged in the circulation of capital does not create surplus value, it is seen by the commercial capitalist as productive because it secures the capitalist's share in existing surplus value and becomes a profit.[56] The emergence of distinct commercial capital enterprises alters the structure of the whole economy and the amount of surplus value available to production capitalists.

Marx then identified 'interest-bearing' capital – capitalists such as banks who earned interest on loans that production capitalists took out to expand production. The generation of interest is possible because, in capitalism, money represents not just purchasing power – buying commodities for consumption – but also the potential to generate more profit in the future through investment as capital.[57] The interest is deducted from the production capitalist's profit rate. Interest-bearing capital, unlike commercial capital, does not lower the general rate of profit; it just subdivides it between recipients of interest and earners of profit.

The relationship between these two types of capital has distinct advantages. It can increase the scale and speed of capitalist production by making it easier to obtain capital and reduce the turnover time (the time it takes for capital to produce, sell and buy new means of production – one 'period' of production). Interest-bearing capital and the credit system it supplies also reduce the importance of commercial capital, for example by shortening the time the production capitalist has to wait for the merchant to return with the proceeds from sales. However, since interest-bearing capital does not produce any surplus value, it is not directly productive.[58]

Finally, in addition to these types of capitalists, Marx identified another: owners of scarce things like land, coal, a patent, a licence to practise law, and so on. Such scarce things can improve productivity above the general

productivity level – the same product can be produced in less labour time or with fewer means of production. That in turn creates ‘surplus profits’ – what Smith and Ricardo might have thought of as ‘rent’ – for capitalists, or landlords and proprietors, who can exploit these advantageous production conditions. Marx thus outlined a theory of ‘monopoly’ gain.

The key, in Marx’s view, is that labour is productive if – and only if – it produces a surplus value for production capital, the engine of the capitalist system; that is, value above and beyond the value of labour power. For Marx, then, the production boundary is defined not by sectors or occupations but by how profits are generated – more specifically, whether an occupation is carried out in a capitalist production context. Only the capitalist enterprise will accumulate the surplus value that can lead to an expansion of production. In this way, the capitalist economy reproduces itself.

Participating in ‘circulation’ or earning interest is not a judgement on such activities’ ‘usefulness’. It was simply necessary, Marx argued, for capital to transform itself from commodity form into money form and back again.[59] In fact, Marx thought that a well-functioning sphere of circulation could raise the profit rate by reducing turnover time for capital. If the ‘circulation sphere’ was not functioning properly – for example, the system of credit that fuelled it was inefficient – it risked absorbing too large a chunk of the surplus value that capitalists hoped to generate by selling their goods and as a result impeding growth.

Marx refined Adam Smith’s distinction between productive (industry) and unproductive (services) sectors into something much more subtle. As can be seen in Figure 6, in Marx’s theory of value every privately organized enterprise that falls within the sphere of *production* is productive, whether it is a service or anything else. Here, Marx’s achievement was to move beyond the simple categorization of occupations and map them onto the landscape of capitalist reproduction.[60] Marx’s production boundary now runs between goods and services production on one side and all those functions of capital that were not creating additional surplus value, such as interest charged by moneylenders or speculative trading in shares and bonds, on the other. Functions lying outside the production boundary take a chunk of surplus value in exchange for circulating capital, providing money or making possible surplus (monopoly) profits.

What is more, in distinguishing between different types of capitalist activity – production, circulation, interest-bearing capital and rent – Marx offers the economist an additional diagnostic tool with which to examine

the state of the economy. Is the sphere of circulation working well enough? Is there enough capacity to bring capital to the market, so that it can be exchanged and realize its value and be reinvested in production? What proportion of profits pays for interest, and is it the same for all capitalists? Do scarce resources, such as 'intellectual' ones like patents on inventions, create advantageous conditions for producers with access to them and generate 'surplus profits' or rents for those producers?

Ricardo and Marx refined the theory of rent to make it clear that rent is income from *redistributing* value and not from creating it. Landlords do not create the soil but they can generate income from their right to exclude from the land others (capitalists) who might use it to produce value. Rent of any kind is basically a claim on the total of social surplus value and therefore lowers productive capitalists' profits. As we will see in the next chapter, neoclassical (mainstream) economics has fundamentally changed this idea of rent into one of imperfections and impediments – which can be competed away.

All these issues have come to the fore again since the 2008 financial crisis. At their heart is how finance has been self-serving, and not actually serving what the American economist Hyman Minsky (1919–96) called the 'capital development of the economy'.[61] In other words, instead of facilitating industrial production, finance has simply degenerated into a casino, aiming to appropriate as much of the existing surplus as possible for itself.[62] But whether that casino is seen as a mere imperfection or as a stable source of unearned income (whereby activities that are not creating value are somehow allowed to be presented a such) makes all the difference in policies that aim to reform the system.

Marx's attempt to define the production boundary was more rigorous than those of Smith and Ricardo and was certainly a long way from those of Petty and King. He introduced the idea of labour power as an objective and invariable standard of value, building on the essential premise shared by earlier economists that value derived from labour. He also shared with them the belief that government was unproductive. The early and classical economists left a legacy of ideas about value – on currencies and protection, free trade, rent, government and technology – which have reverberated down the centuries and remain alive today.

The next chapter explores how, even as the ink was drying on Marx's writing in the British Museum Reading Room, the intellectual world of the classical economists was about to be turned upside down.

2

Value in the Eye of the Beholder: The Rise of the Marginalists

> *… the distribution of the income of society is controlled by a natural law … this law, if it worked without friction, would give to every agent of production the amount of wealth which that agent creates.*
>
> *J. B. Clark,* The Distribution of Wealth: A Theory of Wages, Interest and Profits[1]

In Marx's hands, value theory became a powerful tool for analysing society. While Smith had praised the merits of individual pursuit of happiness and profit, and Ricardo had made the capitalist entrepreneur the hero of the economy, Marx was much more critical of both. As the Industrial Revolution progressed and threw masses of labourers in Europe into urban poverty, his labour theory of value was not just a set of abstract ideas, but an active critique of the system that he saw developing around him. If labour produced value, why was labour continuing to live in poverty and misery? If financiers did not create value, how did they become so rich?

However, the labour theory of value's days were numbered. This chapter is about the emergence of a new set of ideas that inverted the earlier argument that value was nested in objective conditions of production, and that all other economic categories, such as the price of goods and services, were subsumed to it. The classical economists lost their crown to a new dynasty, the neoclassicals.

NEW TIMES, NEW THEORY

Socialist critiques of value theory were multiplying even before Marx wrote *Capital*. A group called the 'Ricardian socialists' used Ricardo's labour theory of value to demand that workers get better wages. They included the Irishman William Thompson (1775–1833), Thomas Hodgskin (1787–1869) and John Gray (1799–1883), both British, and John Bray (1809–97), who was born in the US but worked for part of his life in Britain. Together, they made the obvious argument that if the value of commodities derives from labour, the revenue from their sale should go to workers. This idea underlay the co-operativism of the textile manufacturer Robert Owen (1771–1858), for whom the solution was that workers should also participate in ownership, of both factories and publicly created infrastructure. Marx and Engels were friendly with some of these groups, but very unfriendly towards others whom they thought had no proper analysis of why things were going wrong. The pair collaborated with the groups to whom they were well disposed to produce critiques of capitalism.

Intellectual opposition to capitalism had its practical counterpart in a growing array of radical and socialist political organizations which connected the often dire conditions of working people with programmes of action to remedy them. In Britain, the Chartists (1837–54) demanded reforms to the political system. Trade unionism began to gain a significant following. The Amalgamated Society of Engineers was formed in 1851 and the Trades Union Congress in 1868. During the recession of the 1880s, socialism became more widespread, culminating in the founding of the Labour Party in 1900. Here, Britain was a relative latecomer: the Socialist Workers' Party of Germany was founded in 1875 and the Federation of the Socialist Workers of France four years later.

Faced with these threats to the status quo, the powers that be needed a new theory of value that cast them in a more favourable light. Other influences also encouraged the search for a new analysis of how capitalism works and the troubling question of where value comes from. Malthus's pessimism about the dangers of population growth was an affront to the later-nineteenth-century belief in progress – and the facts did not appear to support him, because the food shortages he predicted had not materialized. Non-conformism offered a moral basis on which to argue that the immiseration of the masses that Marx and others feared was neither inevitable nor desirable. The development of natural sciences and mathematics encouraged attempts to place economics on a similar 'scientific' footing, as opposed to what was becoming seen as the more 'literary' endeavours of the political economists. Above all, perhaps, the

rising power of capitalists in a society long dominated by aristocratic landowners and local gentry meant that a new analysis of capitalism was required to justify their standing.

THE ECLIPSE OF THE CLASSICALS

A series of thinkers and economists who were roughly contemporaneous with Marx began to lay the foundations for what has become modern mainstream economics. Landlords were defended as productive by Lord Lauderdale (1784–1860), a Scottish earl, and profits by Nassau Senior (1790–1864), an English lawyer and economist, as abstinence from consumption. Linking profits to a notion of sacrifice allowed a useful moral justification for the large income inequality between capitalists and workers.[2] Furthermore, as scarce capital could be either invested or saved, profits were no longer linked to theories of exploitation but came to be seen as simply a return for saving and not consuming.

But to put the classicals to bed properly, a new theory of value had to be invented. Two of the principal architects of what became known as neoclassical economics were Léon Walras (1834–1910) and William Stanley Jevons (1835–82). Walras was a professor of economics in Lausanne, Switzerland. For him, 'the characteristic of a science properly speaking is the complete indifference to any consequences, advantageous or undesirable, of its attachment to the pursuit of pure truth'.[3] Walras was keen to show that economics was a real science, less fuzzy than sociology or philosophy, so set out to discover 'pure truths' in the science of theoretical economics rather than focus on applications. Jevons, a Professor of Political Economy at University College, London, began his 1871 *The Theory of Political Economy* with the assertion that economics, 'if it is to be a science at all, must be a mathematical science'. He justified this statement by stating that economics deals with quantities: there were, he continued, 'laws' in economics, which could become like other 'exact' sciences if sufficient commercial statistics were available. Jevons called his economic theory 'the mechanics of utility and self-interest'.

Another economist who linked value to utility was Carl Menger (1840–1921), one of the founders of the 'Austrian school' of economics. As we shall see later, utility is a broad concept, combining ideas about a product's efficiency – is the car reliable? – with vaguer notions of satisfaction and even happiness – does the new car impress the neighbours? For Menger, the value arising from utility set the cost of production; the cost of production, including the cost of labour, did not determine value. Although original, Menger's ideas did not fit comfortably into the new narrative that

economics had to be much more abstract, expressed neatly in mathematical equations based on Newtonian physics.

FROM OBJECTIVE TO SUBJECTIVE: A NEW THEORY OF VALUE BASED ON PREFERENCES

Walras, Jevons and Menger provided a positive and 'scientific' view of reproduction, exchange and income distribution. They used the construct which later came to be called 'marginal utility', and their propagation of a new view on value theory is now referred to as a 'marginal revolution'[4] – it was, however, a slow one.

The marginal utility theory of value states that all income is reward for a productive undertaking. Given the large investments being made in factories and the edifices of the Industrial Revolution, it suited the changing circumstances of the second half of the nineteenth century. But it did not come out of nowhere; indeed, it has a long history. In medieval times, thinkers argued that 'just prices' were those that reflected an object's utility. In his *Summa Theologica*, the thirteenth century philosopher-theologian Thomas Aquinas discussed the concept of the just price in a section of the book called 'Of Cheating, Which Is Committed in Buying and Selling'. Just price was a normative concept, against what was seen as the wrong price resulting from morally evil greed. The medieval Church inveighed against the sin of greed and avarice, which broadly meant profiteering by middlemen and moneylenders. In Dante's *Inferno*, usurers are consigned to the hottest part of hell (circle 7) because they are making money not from the productive sources, which for Dante were Nature or Art, but from speculative changes in interest rates. Indeed, he is so disgusted by usury that he puts usurers just below the circle of hell housing the sodomites.

This normative and moral view of price, linked to cheating or criminal behaviour, began to fade after the seventeenth century – the time of Petty and King – but lingered on until firmly supplanted by the concept of individual utility, which held that it was not about good or bad but how common goals could be reached through each individual trying to maximize the benefit to him- or herself. In 1776 – the year that Adam Smith published *The Wealth of Nations* – the Englishman Jeremy Bentham argued that 'the greatest happiness of the greatest number' should be the 'measure of right and wrong'.[5] In other words, an action should be evaluated according to its consequences in a particular context: killing may be justified if it prevents more killing. This 'utilitarian' theory of ethics spilled over into ideas about production. In France, Jean-Baptiste Say

(1767–1832), Smith's contemporary and a hostile critic of Quesnay, argued in his 1803 book *Treatise on Political Economy* that the value of a commodity resides in its utility to a buyer and, therefore, that productive labour is labour which produces utility. In Say's view, labour in services – which classical economists thought fell squarely into the 'unproductive' category, because they failed to produce 'things' – could in fact be reclassified as productive, so long as those services fetched a price and labour got paid a wage.[6]

The most influential person in developing utility theory was the late-nineteenth-/early-twentieth-century British economist Alfred Marshall (1842–1924), Professor of Political Economy (as it was still called) at Cambridge. Significantly, he was trained as a mathematician. Marshall's 1890 *Principles of Economics* diffused the new ideas to generations of students. The economics library in Cambridge is known simply as the Marshall Library; introductory economics textbooks still include diagrams he developed in the nineteenth century.

In many respects Marshall was a natural heir to the classical tradition. He accepted that the cost of production was important in determining a commodity's value. But he and his followers shifted thinking about value from the study of broad quantities of capital, labour and technology inputs and their returns to that of small incremental quantities. Using mathematical calculus, they focused on how a small – or 'marginal' – change in one variable causes a change in another: for instance, how a small change in price affects the quantity of product demanded or supplied.

So what was the new value theory, marginalism, about? First, it is based on the notions of utility and scarcity and is subjective: the value of things is measured by their usefulness to the consumer. There is, therefore, no 'objective' standard of value, since utility may vary between individuals and at different times. Second, this utility decreases as the amount of a thing that is held or consumed increases. The first Mars Bar you eat in a day may provide a lot of utility or satisfaction and even happiness. It is enjoyable and maybe staves off hunger pangs. But as you go on eating Mars Bars they cease to be so enjoyable and may even make you feel ill. At some point the utility gained from eating them will decrease.[7] In this way, the utility of the last bar is less, possibly much less, than that of earlier bars. This is 'marginal utility' – in the case of a Mars Bar, worth less to you than the previous one, 'decreasing marginal utility'. By the same token, the scarcer a thing is, the more utility it gives you –

‘increasing marginal utility’. One Mars Bar on a desert island can give you more happiness than any number of bars bought from your corner shop.

Prices, then, reflect the utility that buyers get from things. The scarcer they are – the higher their marginal utility – the more consumers will be willing to pay for them. These changes in the marginal utility of a product came to be known as consumer ‘preference’. The same principle applies to producers. ‘Marginal productivity’ is the effect that an extra unit of produced goods would have on the costs of production. The marginal cost of each extra Mars Bar that rolls off the production line is lower than the cost of the previous one.

This concept of marginalism lies at the heart of what is known today as ‘neoclassical’ theory – the set of ideas that followed the classical theory developed by Smith and Ricardo and was extended by Marx. The term *neo*classical reflected how the new theorists stood on the shoulders of giants but then took the theory in new directions. Microeconomic theory, the theory of how firms, workers and consumers make choices, is based on the neoclassical theory of production and consumption which rests on the maximization of profits (firms), and utility (consumers and workers).

As a mathematician, Marshall used mathematical calculus, borrowed from Newtonian physics, to develop his theory of how an economy worked. In his model, the point at which a consumer’s money is worth more to him or her than the additional (marginal) unit of a commodity (that next Mars Bar) that their money would purchase, is where the system is in ‘equilibrium’, an idea reminiscent of Newton’s description of how gravity held the universe together. The smooth, continuous curves of these equilibrating and evolutionary forces depict a system that is peaceful and potentially ‘optimal’. The inclusion of concepts like equilibria in the neoclassical model had the effect of portraying capitalism as a peaceful system driven by self-equilibrating competitive mechanisms – a stark contrast to the ways in which the system was depicted by Marx, as a battle between classes, full of disequilibria and far from optimal, whose resulting revolutions would have been better described by Erwin Schrödinger’s concept of quantum leaps and wave mechanics.

So keen was Marshall to emphasize the equilibrating and evolutionary forces in economics, with their smooth, continuous curves that could be described by mathematical calculus, that the epigraph of his 1890 *Principles of Economics* was the Latin tag *Natura non facit saltum*, a nod to its use by Darwin in his 1859 *On the Origin of Species* to make the point

that Nature, rather than progressing in leaps and bounds, evolves in incremental steps, building on previous changes.

The equilibrium concept had a lot of appeal at the start of the twentieth century, when the rise of socialism and trade unions in Europe threatened the old, often autocratic, order and the conventional wisdom was that capitalism was largely self-regulating and government involvement was unnecessary or even dangerous.

Equilibrium was predicated on the notion of scarcity, and the effect of scarcity on diminishing returns: the more you consume, the less you enjoy each unit of consumption after a certain amount (the maximum enjoyment); and the more you produce, the less you profit from each marginal unit produced (the maximum profit). It is this concept of diminishing returns that allows economists today to draw smooth curves in diagrams, using mathematical calculus, so that maxima and minima points (e.g., the bottom of a U-shaped curve showing how costs change with increased production) provide the equilibrium targets and utility maximization.

Nineteenth-century economists liked to illustrate the importance of scarcity to value by using the water and diamond paradox. Why is water cheap, even though it is necessary for human life, and diamonds are expensive and therefore of high value, even though humans can quite easily get by without them? Marx's labour theory of value – naïvely applied – would argue that diamonds simply take a lot more time and effort to produce. But the new utility theory of value, as the marginalists defined it, explained the difference in price through the *scarcity* of diamonds. Where there is an abundance of water, it is cheap. Where there is a scarcity (as in a desert), its value can become very high. For the marginalists, this scarcity theory of value became the rationale for the price of everything, from diamonds, to water, to workers' wages.

The idea of scarcity became so important to economists that in the early 1930s it prompted one influential British economist, Lionel Robbins (1898–1984), Professor of Economics at the London School of Economics, to define the study of economics itself in terms of scarcity; his description of it as 'the study of the allocation of resources, under conditions of scarcity' is still widely used.[8] The emergence of marginalism was a pivotal moment in the history of economic thought, one that laid the foundations for today's dominant economic theory.

The Marginal Revolution

The ‘marginal revolutionaries’, as they have been called, used marginal utility and scarcity to determine prices and the size of the market. In their view, the supply and demand of scarce resources regulates value expressed in money. Because things exchanged in a monetary market economy have prices, price is ultimately the measure of value. This powerful new theory explained how prices were arrived at and how much of a particular thing was produced.[9] Competition ensures that the ‘marginal utility’ of the last item sold determines the price of that commodity. The size of the market in a particular commodity – that is, the number of items that need to be sold before marginal utility no longer covers the costs of production – is explained by the scarcity, and hence price, of the inputs into production. Price is a direct measure of value.[10] We are, then, a long way from the labour theory of value.

But what this model gains in versatility – the notion that the preferences of millions of individuals determine prices, and hence value – it loses in its ability, or, rather, lack of ability, to measure what Smith called ‘the wealth of nations’, the total production of an economy in terms of value. As value is now merely a relative concept – we can compare the value of two things through their prices and how the prices may change – we can no longer measure the labour that produced the goods in the economy and by this means assess how much wealth was created.

Marginal utility and scarcity need a couple of additional assumptions for price determination to work as intended. First, all humans have to be one-dimensional utility calculators who know what’s best for themselves, what price to pay for what commodity and how to make an economically ‘rational’ choice.[11] Second, there must be no interference, for example by monopolies, in price-setting. ‘Equilibrium’ with ‘perfect competition’ – in which supply and demand are exactly balanced, an idea Jean-Baptiste Say developed back in the early nineteenth century – became a necessary and central concept in economics. These assumptions, as we will see, bear heavily on today’s discussion of value creation.

The Production Boundary Becomes Malleable

The consequences of marginal thinking for the production boundary are dramatic. As we have seen, classical thinkers differed in their definition of who was and was not productive. For Quesnay only farmers were productive; Smith put services in the ‘unproductive’ bracket; and even Marx defined productive workers as those who were working in capitalist production. In marginal thinking, however, such classification was swept

aside. What replaced it was the notion that it is only whatever fetches a price in the market (legally) that can be termed productive activity. Moreover, productivity will fluctuate with prices, because prices determine value, not vice versa. The utility theory therefore completely changes the concept of productive and unproductive labour. In fact, the distinction effectively falls away, since every sector that produces for the market exchanges its products – which means there are now few definitively unproductive sectors. The only part of the economy which clearly lies outside the production boundary and is unproductive, as in Figure 7, consists of those who receive income not earned in the market: the government by collecting taxes and the beneficiaries of government subsidies such as social security payments, and state entities like the armed forces.

In Marshall's state of 'equilibrium', where prices are not distorted, everyone gets paid what they are worth – which may change if consumers alter their tastes or if technology advances. This has important consequences for how incomes are assessed and justified. What workers earn is reflected in their marginal productivity and their revealed preferences (marginal utility) for leisure versus work. There is no longer any room for the analytical distinctions that Ricardo or Marx made about a worker's contribution to production, let alone the exploitation of that worker. You are valuable because what you provide is scarce. Because we are rational utility calculators in the face of scarcity, we don't let things go to waste. Workers might choose unemployment because that gives them more marginal utility than working for that or a given wage. The corollary of this logic is that unemployment is voluntary. Voluntary unemployment arises from viewing economic agents as rationally choosing between work and leisure (i.e. 'intertemporal maximization' in modern theory). In other words, Marx's concept of the 'reserve army of labour' disappears into thin air.

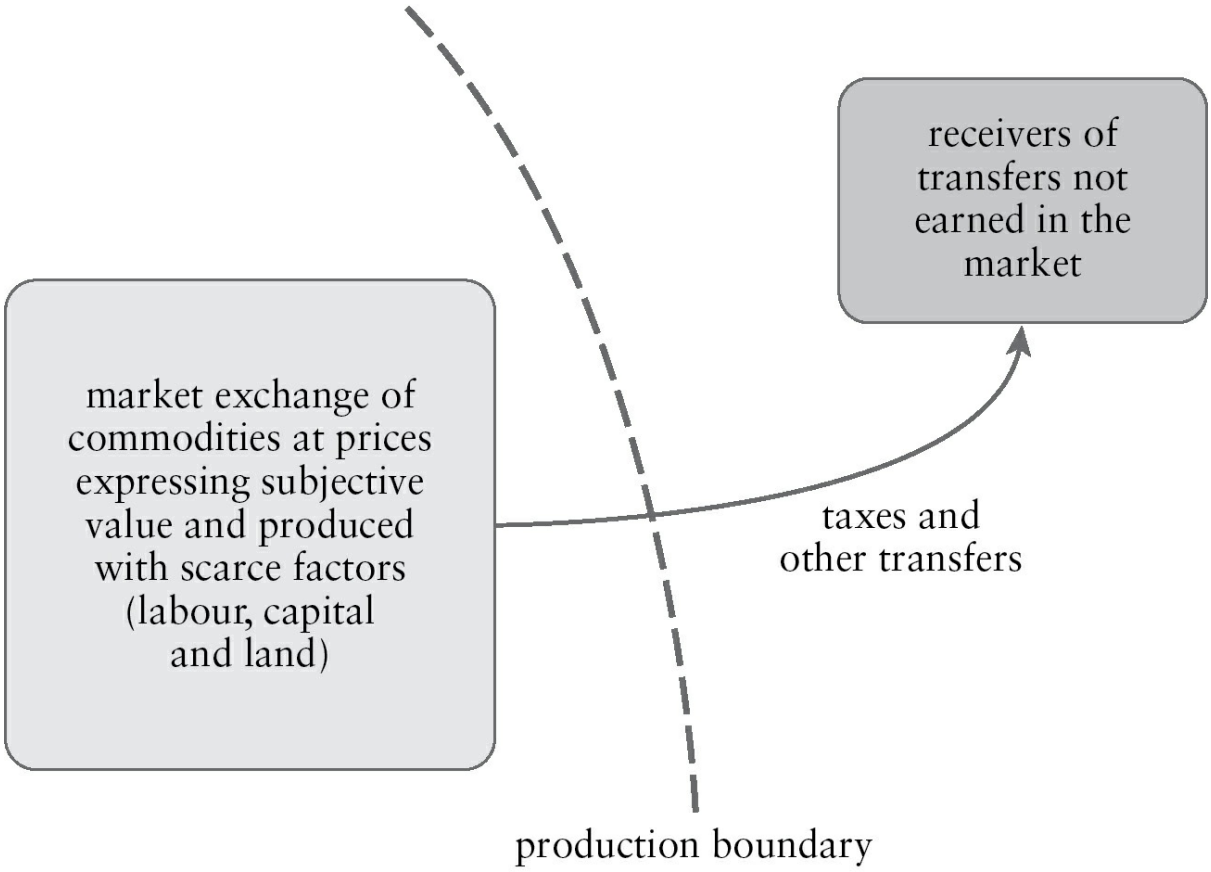


Figure 7. The marginalist revolution

As Lionel Robbins neatly put it,

> In the first place, isolated man wants both real income and leisure. Secondly, he has not enough of either fully to satisfy his want of each. Thirdly, he can spend his time in augmenting his real income or he can spend it in taking more leisure. Fourthly, it may be presumed that, save in most exceptional cases, his want for the different constituents of real income and leisure will be different. Therefore he has to choose. He has to economize.[12]

Inherent in equilibrium is the idea that everything is in everyone's interest. In the 1940s the Russian-born British economist Abba Lerner (1903–82) formulated what he called the 'first fundamental welfare theorem',[13] which basically states that competitive markets lead to 'optimal' outcomes for all. Once market exchange at equilibrium prices has taken place, no one can be made better off, or, in economic parlance, have their 'welfare' increased (for example, by accepting more work) without making someone else worse off.

Today, competitive markets where no one can be made better off without someone being made worse off are known as 'Pareto-optimal' – named after Walras's successor in Lausanne, Vilfredo Pareto (1848–1923), who was the first to introduce the term 'welfare maximization'. In his

Manual of Political Economy (1906), Pareto studied economic equilibrium in terms of solutions to individual problems of 'objectives and constraints', and was the first economist to argue that utility maximization did not need to be cardinal (i.e., the exact amount that someone wanted something) just the ordinal amount (how much they wanted it more than something else – X versus Y). This made mathematical calculus even easier to use, and many welfare properties in economics today bear his name. He used his theories to argue for free trade in Italy, which did not make him popular with the Fascist government of the time, which was more protectionist.

But to get to these 'optimal' outcomes, we must ensure that equilibrium holds: all obstacles to equilibrium, such as an interfering government, monopolies, other rents arising from scarcity and so on, must be obliterated. Our problems, marginalism holds, derive solely from imperfections in, and inhibitions on, the smooth working of the capitalist machine. Rent is no longer seen as 'unearned income', as it was by the classical economists, but as an imperfection that can be competed away. Left to itself, capitalism can thus create maximal value for everyone, which is conveniently what everyone 'deserves' based on their marginal product. The contrast with the classical economists is glaring. For Marx, capitalists appropriate surplus value by paying a wage less than the value of labour. Smith and Ricardo held that value was created by effort that directly added up to the wealth of nations. But with marginal utility there are no longer classes, only individuals, and there is no objective measurement of value.

This approach has a very important consequence. It suggests that government should never intervene in the economy unless there are market failures. Market failure theory uses the first fundamental theorem (FFT) of welfare economics as its starting point. The FFT holds that markets are the most efficient allocators of resources under three specific conditions: first, that there exists a complete set of markets, so that all goods and services which are demanded and supplied are traded at publicly known prices; that all consumers and producers behave competitively; and that an equilibrium exists.

Violations of any of these three assumptions leads to the inefficient allocation of resources by markets, or what marginalists term 'market failures'. Market failures might arise when there are 'positive externalities', benefits to society such as basic science research from which it is hard for individual firms to profit; or 'negative externalities', bad things like pollution, which harm society but are not included in firms' costs. If markets are not 'Pareto-optimal', then everyone could be better

off as a result of public policies that correct the market failure in question.[14] However, as we will see in Chapter 8, a body of economics referred to as Public Choice theory, advocated by Nobel Prize winner James Buchanan (1919–2013), later argued that as government failures are even worse than market failures (due to corruption and capture), so the correction of market failures by bureaucrats might make things even worse.

From the Class Struggle to Profits and Wages in 'Equilibrium'

Defining everything that commands a price as valuable led to the marginalists' conclusion that what you get is what you are worth. Profits are not determined by exploitation but by technology and the 'marginal product of capital'. Capital and labour are seen as the two main inputs into production, and so just as labour earns wages for its productive contribution (marginal product of labour), capital earns a profit (marginal product of capital). John Bates Clark (1847–1938), a former critic of capitalism who converted to become one of the most ardent contributors to the marginalist revolution, argued strongly against the idea that labour was exploited. Capital could not exploit labour, he reasoned, because labour and capital were simply earning their 'just rewards' – their marginal products. In Clark's view, capital goods themselves were the rewards for capitalist self-restraint. Instead of consuming their profits, they had saved them – saving that would eventually result in higher investment in more capital goods (we will come back to this in Chapter 8).

The equilibrium view diverted attention from the tensions between capital and labour, and ultimately from alternative theories on the sources and distribution of value – which almost faded into oblivion from the late nineteenth century onwards, except in expressly Marxist circles and in the thinking of economists such as Joan Robinson (1903–1983), Professor of Economics at Cambridge, and Piero Sraffa (1898–1983), an Italian who also studied and worked in Cambridge. Both were dedicated critics of the neoclassical view of production, believing that the concept of the 'marginal' product of labour and capital was ideologically based, and was also subject to a 'fallacy of composition': the neoclassical theory of production could not apply to the entire system. They engaged actively in what was later called the Cambridge Capital Critique – a debate between the Cambridge, UK-based Robinson and Sraffa, and Solow and Samuelson, who were at MIT in Cambridge, Massachusetts.

Sraffa and Robinson argued that ‘capital’ is heterogeneous and so cannot be used as an aggregate concept. That is, it cannot be aggregated since it would be like adding apples to oranges. In 1952 Robinson, influenced by the writings of Sraffa, argued that the idea of profits as the value measurement of capital is a tautology: there is no way to know the value of capital without knowledge of equilibrium prices, and these require an equilibrium rate of profit that cannot be obtained unless we have estimated the value of capital. Furthermore, following the ideas of Marx, Robinson and Sraffa argued that the rate of profit was not the reward for productive contribution of ‘capital’; it derived from social relations, that’s to say, who owned the means of production and who was forced to work for them. The circularity of the logic of neoclassical theory was partly accepted by Samuelson in a well-known 1966 article in the prestigious *Quarterly Journal of Economics*, where he admitted the logical validity of the points being made by Robinson and Sraffa. Solow, on the other hand, claimed that neoclassical economics should not be distracted by such critiques; and indeed, the debate between the ‘classicals’ and the ‘neoclassicals’ would later disappear, so that most students of economics today don’t even know it happened.

Remarkably, the neoclassical theory of value has not changed much in the last hundred years. The maximization of utility has been extended beyond the economic sphere to explain human behaviour, including crime, drug addiction and, infamously, models of divorce. This particular idea originated with Gary Becker (1930–2014), an American who was Professor of Economics and Sociology at the University of Chicago and won the Nobel Prize in Economics in 1992. In essence, Becker postulated that two individuals marry when there is a positive surplus from their union in contrast to remaining single. These gains may come from, for example, economies of scale, provision of insurance and general risk-sharing. Becker’s ideas encouraged many others to pursue similar investigations.

Attempts have also been made to forge stronger links between macroeconomic patterns (the whole economy, for instance inflation, unemployment and business cycles) and microeconomic decisions made by people and firms. And, as we will see, other work has looked at the need to include non-priced goods (such as care) into GDP.

But despite the critiques, marginal utility theory prevails and is highly influential. The narrow equilibrium view that we will all benefit from perfect competition has influenced – and continues to influence – government policies and those of powerful multilateral bodies such as the

International Monetary Fund and the World Bank: how, with perfect competition, individuals will supposedly maximize their preferences and companies their profits so we will all benefit. On the basis of contemporary economic assumptions, we can no longer reliably say who creates value and who extracts it and therefore how the proceeds of production – income – should reasonably be distributed. In the next chapter we will see how this subjective approach to value has also had a strong impact on the ways we *measure* national wealth and income through the concept of GDP.

THE DISAPPEARANCE OF RENT AND WHY IT MATTERS

When students learn about microeconomics in the classroom (e.g. how prices are determined, including wages), they are not told that this is only one of many different approaches to thinking about value. It is, as far as they are concerned, the only one – and, as a result, there is no need to refer to the word 'value'. The term essentially disappears from the discourse. It is simply Microeconomics 101.

In concluding our history of economic thought, we should ask: is this only an academic exercise, or does it matter? Why it *does* matter is the subject of this book: it is crucial to our understanding of value extraction – and hence the ability to limit it.

The concept of 'rent' has changed in economic thought over the centuries, because rent is the principal means by which value is extracted. The eighteenth-century economists described rent as unearned income, which they thought of as income derived from simply moving existing resources from one hand to another. Their disapproval of unearned income partly came, as we have seen, from medieval strictures on usury – the charging of interest. But it was also practical. Adam Smith believed that a genuinely free market was a market free of rent, and so policymakers had to do their best to eliminate it. His follower David Ricardo considered landowners who collected rent without contributing to the productivity of land to be economic parasites; he denied vehemently that there was any value in the income or rent received from owning land. Rents were *unearned* income and fell squarely outside the production boundary. Both Smith and Ricardo realized that freeing the economy from rent called for strong intervention – in practice by government – to prevent value extraction. Neoclassical economists too; they see rent as an impediment to 'free competition' (free entry and exit of different types of producers and consumers). Once those impediments are removed, competition will benefit everyone.

In the subjective marginalist's approach, wages, profits and rent, along with wages and profits, all arise from 'maximizing': individuals maximizing utility and firms maximizing profits. Thus labour, capital and land are input factors on the same footing. The distinction between social classes, including who owns what, is obliterated, since whether one lends out capital or works for wages depends on an unexplained initial endowment of resources.[15]

Wages are determined by the worker equalizing the (diminishing) marginal utility of the money obtained from working with the 'disutility' of working, for example less leisure time. At the prevailing wage rate, the amount of time spent on work determines the income. This assumes that the amount of employment can be flexibly adjusted. If this is not the case, the marginal utility of taking a job might become less than the utility derived from an equivalent time of leisure; someone chooses not to work. As we have seen, this means that unemployment is therefore voluntary.

Profits and rent are thus determined analogously: the owners of capital (money) will lend it until the marginal utility from doing so is lower than that of consuming their capital. Landlords do the same with their land. For instance, the owner of a house might rent it out and then decide to let her daughter live there for nothing, effectively consuming capital because rent earnings are forgone. The justification for any profits is thus related to individual choices (based on psychology) and the *psychological assumption* that people derive less utility from future consumption (discounting). So the return on capital and land is seen as compensation for future marginal utility at a level which could be enjoyed today if the capital were consumed instead of lent.

In classical economics, therefore, rents are part of the 'normal' process of reproduction. In neoclassical economics, rents are an equilibrium below that which is theoretically possible – 'abnormal' profits. The main similarity is that both theories see rent as a type of monopoly income. But rent has a very different status in the two approaches. Why? Chiefly because of the divergent value theories: classical economics fairly clearly defines rent as income from non-produced scarce assets. This includes, for example, patents on new technologies which – once produced – need not be reproduced any more; the right to issue credit money, which is restricted to organizations with a banking licence; and the right to represent clients in court, which is restricted to members of a Bar association.[16] Essentially, it is a claim on what Marx called the pool of social surplus value – which is enormous compared to any individual

production capitalist, circulation capitalist, landowner, patent holder and so on.

By contrast, in neoclassical economics – in general equilibrium – incomes must by definition reflect productivity. There is no space for rents, in the sense of people getting something for nothing. Tellingly, Walras wrote that the entrepreneur neither adds nor subtracts from value produced.[17] General equilibrium is static; neither rents nor innovation are allowed. A relatively recent refinement, the more flexible partial equilibrium analysis, allows us to disregard interactions with other sectors and introduce quasi-rents, and has since the 1970s led to the idea of 'rent-seeking' by creating artificial monopolies, for example tariffs on trade. The problem is that there is no hard-and-fast criterion with which to assess whether the entrepreneur creates 'good' new things or is imposing artificial barriers in order to seek rents.

The neoclassical approach to rent, which largely prevails today, lies at the heart of the rest of this book. If value derives from price, as neoclassical theory holds, income from rent must be productive. Today, the concept of *unearned income* has therefore disappeared. From being seen by Smith, Ricardo and their successors as semi-parasitic behaviour – extracting value from value-creating activity – it has in mainstream economic discourse become just a 'barrier' on the way to 'perfect competition'. Banks which are judged 'too big to fail' and therefore enjoy implicit government subsidy – a form of monopoly – contribute to GDP, as do the high earnings of their executives.

Our understanding of rent and value profoundly affects how we measure GDP, how we view finance and the 'financialization' of the economy, how we treat innovation, how we see government's role in the economy, and how we can steer the economy in a direction that is propelled by more investment and innovation, sustainable and inclusive. We begin by exploring in the next chapter what goes into – and what is omitted from – that totemic category, GDP, and the consequences of this selection for our assessment of value.

3

Measuring the Wealth of Nations

What we measure affects what we do; and if our measurements are flawed, decisions may be distorted.
Joseph Stiglitz, Amartya Sen and Jean-Paul Fitoussi, *Mismeasuring Our Lives* (2010)

Scarcely a day goes by without politicians, the media or experts opining on the state of a country's GDP – the measure used to calculate the growth of goods and services in an economy: the 'wealth of nations'. Success or failure – real or imagined – in managing GDP can make or break governments and careers. If GDP falls for more than two consecutive quarters, there are cries of 'recession'. If the fall is sustained over a year, it's a depression. But where does this measure come from? And how is it influenced by the way value is understood?

Marginal utility is today a major influence on the measurement of economic activity and growth. It has an effect on the rationale for the kinds of economic activities that are considered productive – which, as we saw in Chapter 2, is basically anything that fetches a (legal) price in the market. According to marginalists, because value derives from price, somebody earning a very high salary is indicative of their productivity and worth. At the same time, anybody holding down a job at all is supposed to reflect their preference for work: the utility of work against that of leisure. GDP can be measured as the total amount of products produced, the total amount demanded, or the total income earned (with adjustments reviewed below). But if income is not necessarily a sign of productivity but of something else – for example the classicals' notion of rent as 'unearned income' – what are the implications for GDP as a reliable measure of an economy's productiveness?

A rise in incomes in the financial sector, for example, would have an effect on GDP. So how sectors are valued influences our calculations of growth rates, and this may in turn influence how we decide to steer the economy. In other words, how we measure GDP is determined by how we value things, and the resulting GDP figure may determine how much of a thing we decide to produce. Performativity!

But if there are problems with the way in which we measure GDP, policymakers can receive misleading signals about what is productive and how to steer the economy. Discussion about which parts of society are productive and which non-productive has been much less explicit since the arrival of marginal utility theory. As long as products and services fetch a price on the market, they are worthy of being included in GDP; whether they contribute to value or extract it is ignored. The result is that the distinction between profits and rents is confused and value extraction (rent) can masquerade as value creation.

This chapter will look at the ways in which governments have calculated growth through national accounting methods, the relationship between these methods and value theory, and the very strange results that have ensued, including the undervaluation of certain activities (like caring for our children), and the overvaluation of others (such as polluting businesses). In Chapter 4 we will see how marginal utility theory has also failed to account for one of the key problems in modern capitalism: the extractive activities of the financial sector.

GDP: A SOCIAL CONVENTION

It is crucial to remember that all types of accounting methods are evolving social conventions, defined not by physical laws and definite ‘realities’ but reflecting the ideas, theories and ideologies of the age in which they are devised.[1] The way in which a spreadsheet is constructed in itself reflects values. An interesting example is the Jesuit Order. Back in the 1500s, the newly founded Order devised an innovative accounting system which blended vision with finance. In order to align finance with the values of their order, they made sure that the cash box could only be opened with two keys: one operated by the person in charge of the finances (the procurator, today’s CFO) and another by the person in charge of the strategy (the rector, today’s CEO).[2] As this instance shows, accounting is not neutral, nor is it set in stone; it can be moulded to fit the purpose of an organization and in so doing affect that organization’s evolution.

In this same way, the modern accounting concept of GDP is affected by the underlying theory of value that is used to calculate it. GDP is based on

the ‘value added’ of a national economy’s industries. Value added is the monetary value of what those industries produce, minus the costs of material inputs or ‘intermediate consumption’: basically, revenue minus material input cost. Accountants call the intermediate inputs a ‘balancing’ item because they balance the production account: cost and value added equal the value of production. Value added, however, is a figure specifically calculated for national accounting: the residual difference (residual) between the resource side (output) and the use side (consumption).

The sum of all industry value-added residuals in the economy leads to ‘gross value added’, a figure equal to GDP, with some minor corrections for taxes. GDP can be calculated either through the production side, or through the income side, the latter by adding up the incomes paid in all the value-adding industries: all profits, rents, interest and royalties. As we will see below, there is a third way to calculate GDP: by adding up expenditure (demand) on final goods, whose price is equal to the sum of the value added along the entire production chain. So GDP can be looked at through production (all goods and services produced), income (all incomes generated), or demand (all goods and services consumed, including those in inventory).

So which industries add value? Following marginalist thinking, the national accounts today include in GDP all goods and services that fetch a price in the market. This is known as the ‘comprehensive boundary’. As we saw in Chapter 2, according to marginalism the only economic sectors outside the production boundary are government – which depends on taxes paid by the productive sectors – and most recipients of welfare, which is financed from taxation. Adopting this principle to calculate GDP might seem logical. But in fact it throws up some real oddities which call into question the rigour of the national accounting system and the way in which value is allocated across the economy. These oddities include how government services are valued; how investments in future capacity, such as R&D, are measured; how jobs earning high incomes, as in the financial sector, are treated; and how important services with no price (such as care) or no legal price (such as the black market) are dealt with. In order to explain how these oddities have arisen, and why the system seems to be so idiosyncratic, we need briefly to look at the way in which national accounting and the idea of ‘value added’ has developed over the centuries.

A Brief History of National Accounts

Value theory has been at the heart of national accounting for a very long time. The most significant early initiatives took place in late-eighteenth-century France, when there were at least eight attempts by different thinkers to estimate France's national product based on Quesnay's land theory of value. Because, as we have seen, for Quesnay the production boundary encompassed only agricultural output – everyone else being classed as living off transfers from the agricultural sector – manufacturing was placed on the unproductive side of the boundary, in the process ignoring dissenting voices such as that of Say, who, taking a broadly utilitarian approach, argued that productive labour is simply labour which produces utility. If the product is something people want to buy – has utility for them – then making it is productive.

Excluding manufacturing from the national product seemed as obvious to Quesnay's followers as it is for us today to include everything that fetches a price. These early French estimators were illustrious figures. They included the writer Voltaire (1694–1778); Antoine Laurent Lavoisier (1743–94), one of the founders of modern chemistry; and his friend the mathematician Joseph Louis Lagrange (1736–1813), better known today for his work on mechanics and mathematical techniques which is still used by economists. Quesnay's ideas proved remarkably durable: as late as 1878, one French estimate of national product was based on his reasoning.[3]

Similarly influential were Adam Smith's ideas of value production. His national income estimates defined only the production or income of agricultural and industrial labour, which produced material goods – actual stuff – and excluded all services, whether government or banking ones. Smith's ideas even underpinned the first account of national product in revolutionary France when, in 1789, Napoleon commissioned Smith's disciple Charles Ganilh (1758–1836) to provide an up-to-date and accurate picture of French national income.[4]

In the late nineteenth century, marginal utility theory predominated. Although radically different from the thinking of earlier economists, it continued to underscore the importance of value theory in national accounting. Increasingly, under its influence, national accountants included everything bought with income: for them, the sum of revenue from market activity, irrespective of sector, added up to the national income. As income tax statistics became more readily available, it was easier to construct estimates based on income data and to analyse the personal distribution of income.

Alfred Marshall, the father of marginal utility theory in Britain, was the driving force behind its application to national income estimates.[5] In his

highly influential *Principles of Economics* he wrote explicitly about how the national product could be estimated. An earlier book, *The Economics of Industry*, co-authored with his wife Mary Paley Marshall (1850–1944), was clear on the utility basis of national income: 'everything that is produced, in the course of a year, every service rendered, every fresh utility brought about is a part of the national income'.[6]

Meanwhile, the labour theory of value which, fully developed by Marx, rooted productivity firmly in the concept of the production of 'surplus value', was either disputed or, increasingly, ignored altogether in assessments of national income. By the early twentieth century it was associated with a revolutionary programme and therefore could not, by definition, sit easily with official statistics in the very nations of which Marxists were so critical. Things were of course different in the countries where Communists came to power: first in the Soviet Union after the 1917 Bolshevik Revolution and later in Eastern Europe after the Second World War (though in justifying their construction of a 'material product system' that valued only material goods they should technically have been invoking Smith, not Marx). With the exception of these socialist states, the idea that assessments of national income should be based on the sum total of all incomes, thus forming a 'comprehensive' production boundary, spread rapidly to many countries.[7]

In the first half of the twentieth century marginalists had become aware of their theory's limitations, and began to debate the inclusion of non-market activities in national income accounting. One of Alfred Marshall's students, the British economist Arthur Cecil Pigou (1877–1959), who succeeded him as Professor of Political Economy at Cambridge, argued that since market prices merely indicated the satisfaction (utility) gained from exchange, national income should in fact go further: it should measure welfare. Welfare, Pigou argued, is a measure of the utility that people can gain through money – in other words, the material standard of living. In his influential 1920 book *The Economics of Welfare*, Pigou further defined 'the range of our inquiry' as being 'restricted to that part of social welfare that can be brought directly or indirectly into relation with the measuring-rod of money'.[8] On the one hand, Pigou was saying that all activities which do not really improve welfare (recall the discussion of welfare principles from Pareto discussed in Chapter 2), should be excluded from national income, even if they cost money. On the other hand, he stressed, activities which do generate welfare should be included – even if they are not paid for. In these, he included free or subsidized government services.

One of Pigou's most prominent disciples was the first person to provide an estimate of the fall in national income of the United States during the Great Depression. The Belarusian-born Simon Kuznets (1901–85), a Professor of Economics at Harvard, won the Nobel Prize in Economics in 1971 for his work on national accounts. Believing that they incurred costs without adding to final economic output, Kuznets, unlike Pigou, excluded from the production boundary all government activities that did not immediately result in a flow of goods or services to households – public administration, defence, justice, international relations, provision of infrastructure and so on.[9]

Kuznets also believed that some household expenditure did not increase the material standard of living, but simply paid for the cost of modern life – in particular the 'inflated costs of urban civilization', such as having to maintain a bank account, pay trade union dues or the social obligation to be a member of a club. Kuznets estimated that between 20 and 30 per cent of consumer expenditure went on such services.[10] However, he did argue that unpaid housework should be included, because it clearly improves economic welfare. Kuznets, then, drew the production boundary according to what he believed improved the material standard of living and what did not.

Perhaps Kuznets's view would have had more traction in a peaceful world. But the exigencies of the Second World War, which forced governments to focus on the war effort, took economists down a different path: estimating output rather than concerning themselves with welfare. As a result, economists who believed that national product is the sum total of market prices prevailed.

The resulting ways in which estimates of output were used to calculate GDP appeared to follow marginal utility theory, but were in fact doubly out of sync with it. First, they ignored the idea of value as utility – the benefit it provided to the consumer – and included items that the Pigou–Kuznets welfare concept would have seen as 'necessary' for the creation of value. Rather than assess whether final consumption increased utility, they added *any* final consumption to national income. In Kuznets's own words: 'Many foods and drugs are worthless by scientific standards of nutrition and medication; many household appurtenances are irrelevant to any scientifically established needs for shelter and comfort; many service activities as well as commodities are desired for the sake of impressing foreigners or our fellow countrymen and could hardly measure up to ethical principles of behaviour in relation to the rest of mankind.'[11] From that perspective, the new national accounts overstated welfare.

Second, competition in economies is generally imperfect – a reality that has proven distinctively uncomfortable for national accountants trained in the neoclassical ideas of perfect competition and ‘equilibria’. By simply adding up market prices they ignored the fact that those prices would not always produce an equilibrium and be compatible with ‘perfect competition’; prices could therefore be higher or lower than if equilibrium prevailed, thereby giving a distorted impression of value creation. In short, during the war years practice became significantly detached from the prevailing theory – or, seen another way, the utility theory of value did not solve the urgent war-related problems of the time.

In many ways, the national accounts as we know them today stem from the trauma of the Great Depression of the 1930s, and the needs of the Second World War war effort. In this, as in so much else, the British economist John Maynard Keynes (1883–1946) was a pivotal figure. In his 1936 masterpiece *The General Theory of Employment, Interest and Money*, written during the Great Depression, Keynes assumed that workers would underestimate the purchasing power of their wages, and would therefore be willing to produce more than they needed to. In this way, workers’ involuntary overproduction would in turn create involuntary unemployment – fewer workers being needed to do the same amount of work – and the economy could find itself in a low-output equilibrium. This is a state in which forces in the economy, such as supply and demand, are in balance and there is no incentive to change them, even though the total output of the economy is low and wages and employment are depressed. Keynes used this idea to develop a theory of the macroeconomy – the economy as a whole – in which government spending could stabilize the business cycle when business was investing too little, and even raise the economy’s output.

In order to lift the economy out of the depression, governments needed information to measure how their policies were working. Up until then, they had flown mostly blind: they had no need for detailed statistics because the economy was supposed to be self-regulating. Keynes’s book *How to Pay for the War*, published in 1940, introduced the idea of recording national income in a set of accounts and completely changed the way in which governments used that data.

In the late 1930s and the 1940s national accountants took up Keynes’s ideas about how government could invigorate an economy, and came to view government spending as directly increasing output. For the first time in the history of modern economic thought, government spending became important – in stark contrast to Kuznets’s omission of many government

services from the national income. This redefinition of government as a contributor to national product was a decisive development in value theory. Keynes's ideas quickly gained acceptance and were among the main influences behind the publication of the first handbook to calculate GDP, the United Nations' System of National Accounts (SNA): a monumental work that in its fourth edition now runs to 662 pages.

THE SYSTEM OF NATIONAL ACCOUNTS COMES INTO BEING

After the Second World War, formal international rules were drawn up, standardizing national accounting for production, income and expenditure. The first version of the SNA, compiled by the United Nations, appeared in 1953.[12] The SNA describes itself as 'a statistical framework that provides a comprehensive, consistent and flexible set of macroeconomic accounts for policymaking, analysis and research purposes'.[13] It defines national accounting as measuring 'what takes place in the economy, between which agents, and for what purpose'; at its heart 'is the production of goods and services'.[14] GDP is '[i]n simple terms, the amount of value added generated by production'.[15] It is defined explicitly as a measure of value creation. It can therefore be said that the national accounts, too, have a production boundary.

The SNA's emergence in the early post-war years owed much to recent economic, political and intellectual developments. The experience of depression and war weighed heavily on policymakers' minds. Many countries saw wartime planning, which was based on unprecedented amounts of economic information, as a success. Political pressures were important too. In the US, the New Deal of the 1930s and full employment during the war led many voters to believe that government could intervene benignly and progressively in the economy. In Europe, the strength of left-wing parties after the war – exemplified by the Labour Party's 1945 election victory in the UK – also changed, and marked a change in, people's attitudes, and made fuller and more accurate national accounts essential. The crucial question was, and remains: on what theory of value were they based?

'Simple' national income estimates had to add up the price of production (minus intermediate goods) in the economy, or incomes, or the expenditure of all economic actors on final goods: *National Production = National Income = National Expenditure*. In order to carry out this estimate, we might have expected the SNA's authors to use as their methodology the prevailing economic theory of value, marginal utility. But they didn't – or, at least, not fully. In fact, the resulting model was,

and is, a strange muddle in which utility is the major, but not the only, ingredient.

The SNA brings together various different ways of assessing the national income that had developed over centuries of economic thinking. Decisions about what gets included in the production boundary have been described as 'ad hoc',[16] while national accountants admit that the SNA rules on production are 'a mix of convention, judgment about data adequacy, and consensus about economic theory'.[17] These include devising solutions based on 'common sense'; making assumptions in the name of 'computational convenience' – which has important consequences for the actual numbers we come up with when assessing economic growth; and lobbying by particular economic interests.

In fairness, there have always been practical reasons for this ad hoc approach. Aspects of the economy, from R&D and housework to the environment and the black economy, proved difficult to assess using marginal utility. It was clear that a comprehensive national accounting system would have to include incomes from both market exchange *and* non-market exchange – in particular, government. With market-mediated activity lying at the heart of the marginalist concept of value, most estimators of national income wanted to adopt a broader approach.[18]

National Income Accounting Gets it All Together

National income accounting, then, incorporates many different accounting methods. The system simultaneously allows an integrated view of the different aspects of the economy – both production (output) and distribution of income – and obliges national income accountants to link each commodity produced with someone's income, thereby ensuring consistency. To maintain consistency between production, income and expenditure, the national accounts must record all value produced, income received, money paid for intermediate and finished goods and so on as transactions between actors in the economy – the government, or households or a particular sector – in which each actor has an account. This helps provide a detailed picture of the economy as a whole.

Expenditure on final goods must add up to GDP (as the price of intermediate goods goes into the price of the final goods). So it is possible to compute GDP from the expenditure side and, as we saw in Chapter 1, Petty used this method to estimate national output as early as the seventeenth century. Modern national accounts divide expenditure into the following categories:

GDP = Consumption by households (C) + Investment by companies and by residential investment in housing (I) + Government spending (G).

This can be expressed as: GDP = C + I + G.

For simplicity's sake we will ignore the contribution of net exports. Two observations are in order: first, on the expenditure side, companies only appear as investors (demanding final investment goods from other companies). The remainder of spending (aggregate demand) is split between households and the government. Government expenditure is only what it spends itself; that is, excluding the transfers it makes to households (such as pensions or unemployment benefits). It is its collective consumption expenditure on behalf of the community. By focusing on government only in terms of the spending, it is by definition assumed to be 'unproductive' – outside the production boundary.

MEASURING GOVERNMENT VALUE ADDED IN GDP

In Chapter 8 we will see that government is rarely acknowledged as a creator of value – indeed, quite the reverse. Yet national accounting conventions have in fact been quietly tracking its value-added contribution for the last half century—and it's not small! While Chapter 8 is fully dedicated to looking at the various ways in which economists, and those they advise, have considered government – as a value-creating entity or just a strain on the economy – in this chapter we focus on the relevance of this discussion to how GDP is calculated. The most striking thing to emerge from these analyses is that, contrary to the views of most economists, government certainly does add value to the economy,

Figure 8 below tracks US government value added and expenditure since 1930. During the Second World War the government bought an astonishing half of national output. In this figure we can see that government value added has hovered between 11 and 15 per cent of GDP for the post-war period. As a comparison, the finance industry adds some 4 per cent of GDP in the US and 8 per cent in the UK. But the chart reveals what looks like a strange discrepancy. It shows that government expenditure has been consistently higher than government value added, at between 20 and 25 per cent of GDP.

It is important to stress, however, that the difference between value added and final expenditure is not the government's budget deficit. Rather, the deficit is government revenue (mainly taxes) minus expenses, including transfers of funds from the government to households, such as pensions and unemployment benefit – which, since households spend the money from pensions and benefits, are defined in national accounting as

household, rather than government, spending (it's the final expenditure that matters, remember). It is that household spending that counts towards final demand for the whole economy. So, what is going on?

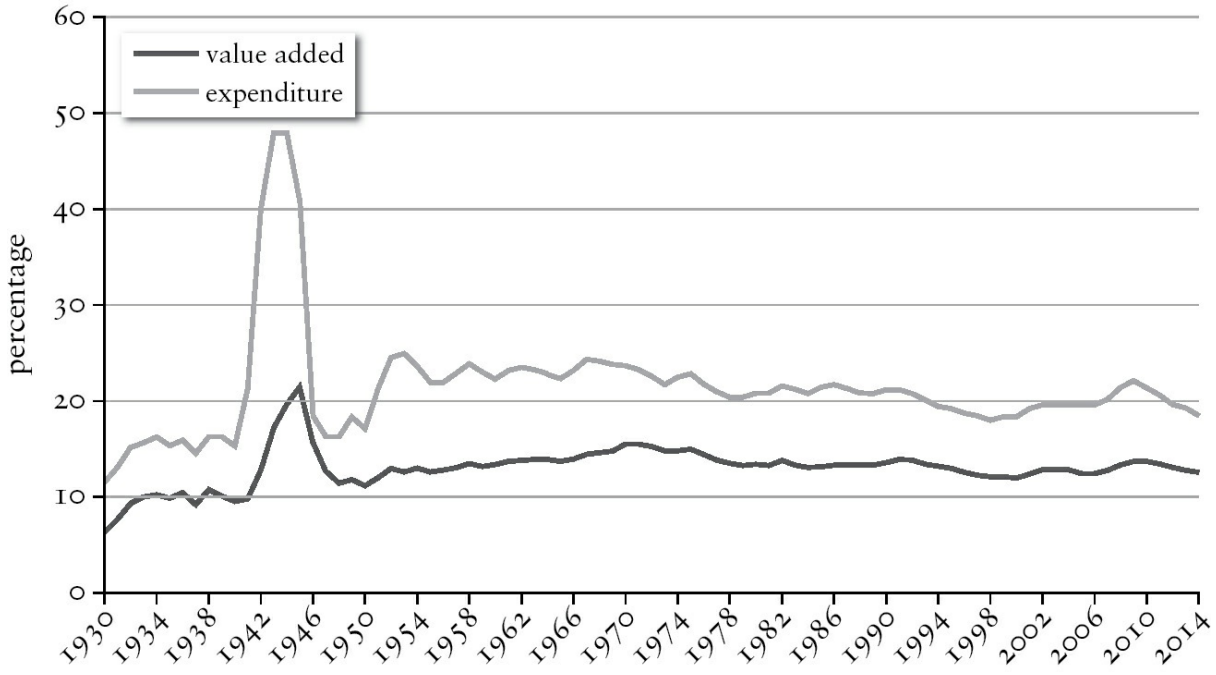


Figure 8. US government expenditure and value added as a share of GDP, 1930–2014[19]

Spending and Value

Before answering this question, we have to recognize that government value added cannot be computed in the same way as that of other industries and, as a result, is a complicated issue for national statistical institutes. A lot of government activities are not sold at market prices: that is, prices that pay for all production costs (including wages, rent, interest, royalties as well as inputs into production) and yield a profit for a private-sector business. Instead, government activities are provided at lower, 'non-market' prices – or even for free. Consider schools, state-funded universities, public healthcare, public transport, parks, recreation and the arts, police and fire services, the law courts, environmental protection such as flood prevention and so on. These goods are largely financed by taxes or debt.

Given these lower prices, the usual way of calculating value added for a business doesn't work with government activities. Let's recall that value added is normally the value of output minus costs of intermediate inputs used in production. The value added by a business is basically workers' wages plus the business's operating surplus, the latter broadly similar to

gross operating profit in business accounting terms. So adding up the non-market prices of government activities is likely to show less value added, because they are set with a different, non-commercial objective: to provide a service to the public. If the non-market prices of the output are lower than the total costs of intermediate inputs, value added would even show up as negative – indeed, government activities would 'subtract' value. However, it makes no sense to say that teachers, nurses, policewomen, firefighters and so on destroy value in the economy. Clearly, a different measurement is needed. As the British economist Charles Bean, a former Deputy Governor for Economic Policy at the Bank of England, argues in his *Independent Review of UK Economic Statistics* (2016),[20] the contribution to the economy by public-sector services has to be measured in terms of 'delivering value'.[21] But if this value is not profit, what is it?

National accountants have therefore long adopted the so-called 'inputs = outputs' approach. Once the output is defined, value added can be computed because the costs of intermediate inputs, such as the computers that employees use, are known. But since government's output is basically intermediate inputs plus labour costs, its value added is simply equal to its employees' salaries. One significant consequence of this is that the estimate of government value added – unlike that of businesses – assumes no 'profit' or operating surplus on top of wages. (In Figure 8 above, the dark-grey line shows the value added of government; it is equal – with slight adjustments – to the share of government employment income in GDP.) In a capitalist system in which earning a profit is deemed the outcome of being productive, this is important because it makes government, whose activities tend to be non-profit, seem unproductive.

But then what about the light-grey line, which represents government final consumption expenditure? We have already seen that pensions and unemployment benefits paid by government are part of household final consumption, not government spending. More broadly, it is not obvious why government should have *any* final consumption expenditure in the way that households do. After all, companies are not classed as being final consumers; their consumption is seen as intermediate, on the way to producing final goods for households. So why isn't government spending likewise classed as intermediate expenditure? After all, there are, for example, billions of school students or medical patients who are consumers of government services.

Indeed, following this logic, government is also a producer of intermediate inputs for businesses. Surely education, roads, or the police, or courts of law can be seen as necessary inputs into the production of a

variety of goods? But herein lies is a twist. If government spending were to increase, this would mean that government was producing more intermediate goods. Businesses would buy at least some of those goods (e.g. some public services cost money) with a fee; but because they were spending more on them (than if government was not producing anything, and therefore not buying supplies from businesses), their operating surplus and value added would inevitably fall. Government's share of GDP would rise, but the absolute size of GDP would stay the same. This does, of course, run counter to Keynesian attempts to show how increases in government demand could lift GDP.

Many economists made exactly this argument in the 1930s and 1940s – in particular Simon Kuznets, who suggested that only government non-market and free goods provided to households should be allowed to increase GDP. Nevertheless, the convention that all government spending counts as final consumption arose during the Great Depression and the Second World War, when the US needed to justify its enormous government spending (the spike in the light-grey line in Figure 8 in the early 1940s). The spending was presented as adding to GDP, and the national accounts were modified accordingly.[22]

Later in the twentieth century there were repeated attempts to clear up the confusion over whether certain kinds of government spending counted as intermediate or final consumption. This was done by identifying which government activities provided non-market and free services for households (for example, schools), as opposed to intermediate services for businesses (for example, banking regulation). The distinction is not easy to make. Governments build roads. But how much of their value accrues to families going on holiday and how much to a trucking company moving essential spare parts from factory to user? Neither family nor trucking company can build the road. But the family on holiday adds to total final demand; the trucking company is an intermediate cost for businesses.

In 1982, national accountants estimated that some 3 to 4 per cent of Swedish, German and UK GDP was government expenditure that, previously categorized as final consumption expenditure, should be reclassified as (intermediate) inputs for businesses. This had the effect of lowering government's overall value added by between 15 to 20 per cent.[23] To take an example of such reclassification, in 2017 the UK telecommunications regulator, Ofcom, compelled British Telecom (a private firm) to turn its broadband network operation Openreach into a separate company following repeated complaints from customers and other broadband providers that progress in rolling out broadband around the

country had been too slow and that the service was inadequate. At least part of the cost of Ofcom could be seen as beneficial to the private telecommunications sector. Yet the convention that all government spending should count as final consumption has proved remarkably resistant to change.

Now we can see why government final consumption expenditure is bigger than its value added in Figure 8. The government's value added only includes salaries. However, the government also purchases a lot of goods and services from businesses, from coffee to cars, from pencils to plane tickets, to the office rentals for regulating bodies such as Ofcom. The producers of these goods and services, not the government, take credit for the value added. Since government is treated as a final consumer, the purchase of goods and services increases its spending. Clearly, government expenditure can be higher than what it charges (e.g. fees for services) because it raises taxes to cover the difference. But need the value of government be undermined because of the way prices are set? By not having a way to capture the production of value created by government – and by focusing more on its 'spending' role – the national accounts contribute to the myth that government is only facilitating the creation of value rather than being a lead player. As we will see in Chapter 8, this in turn affects how we view government, how it behaves and how it can get 'captured' easily by those who confidently see themselves as wealth creators.

SOMETHING ODD ABOUT THE NATIONAL ACCOUNTS: GDP FACIT SALTUS!

Apart from this curious view of government, the national accounts expose a number of other accounting oddities. GDP, for instance, does not clearly distinguish a cost from an investment in future capacity, such as R&D; services valuable to the economy such as 'care' may be exchanged without any payment, making them invisible to GDP calculators; likewise, illegal black-market activities may constitute a large part of an economy. A resource that is destroyed by pollution may not be counted as a subtraction from GDP – but when pollution is cleaned up by marketed services, GDP increases. And then there's the biggest oddity of all: the financial sector.

Does the financial sector simply facilitate the exchange of existing value, or does it create new value? As we will see in Chapters 4 and 5, this is the billion-dollar question: if it's answered wrongly, it may be that the growing size of the financial sector reflects not an increase of growth, but rent being captured by some actors in the economy. First, however, there are some other inconsistencies to be considered.

Investment in Future Capacity

First, let's look at how R&D is dealt with in the national accounts. Before 2008, the SNA considered in-house R&D to be an input into production[24] – in other words, a company's spending on R&D (research equipment, laboratories, staff and the like) was treated as a cost and subtracted from the company's final output. However, in the 2008 version of the SNA, in-house R&D was reclassified as an investment in the company's stock of knowledge, to be valued 'on the basis of the total production costs including the costs of fixed assets used in production'.[25] It became a final productive activity rather than just an intermediate cost towards that activity.

The SNA's decision to reclassify R&D was justified less by value theory than by 'common-sense' reasoning: the contribution of 'knowledge' to production seemed to be significant, and should therefore be recognized. R&D was made productive because it was considered important.

As a result, since 2008 GDP has been enlarged by the annual cost of R&D, including the depreciation of fixed assets used. When in 2013 the US implemented this change, the value from R&D added $400 billion – 2.5 per cent of US GDP – to national income overnight.[26] Of course, those sectors with the largest R&D contributions improved their share of GDP, making them look more important than others.

The Value of Housework … and the House

Then there's housework. Feminists in particular have long objected to the lack of recognition given to housework's contribution to the economy. The national accounts exclude all housework, and therefore a large part of women's work, from production. The architect of the first and second editions of the SNA (1953 and 1968), the Nobel Prize-winning British economist Sir Richard Stone (1913–91) – sometimes called 'the father of national income accounting' – had decided views on the matter. Writing for the UN committee that drafted the first SNA evaluation of household production, Stone commented that it 'is unnecessary to impute an income to family services or to the services of household equipment and may even prove an embarrassment to do so, since, not only are there very little data in this field, but the principles on which such imputations should be made are obscure'.[27] He simply thought it was impossible to know how to do it – and even if a solution could be devised, doing so would be socially awkward.

Now, seventy years later, since there is still no theory – beyond ignorance or shame – that explains why housewives (and house husbands) should not be included in GDP, the SNA architects have come up with a different defence. They have expressed a 'reluctance' to include such work because, although it is equivalent to work done by servants, 'By convention … only the wages of the domestic staff are treated as the value of output.'[28] The 'convention' here is ironically close to Marx's value theory that only someone who produces a surplus for a capitalist generates surplus value. But Marx's point was linked to his value theory and understanding of how capitalism works (or does not work), whereas in this instance the convention has been cherry-picked because it is convenient for the current system.

In explaining why housework is accounted as unproductive, national accountants are forced constantly to fall back on their 'comprehensive' production boundary, and are at pains to invoke 'common sense'. Their explanations include: 'the relative isolation and independence of these activities from markets, the extreme difficulty of making economically meaningful estimates of their values, and the adverse effects it would have on the usefulness of the accounts for policy purposes and the analysis of markets and market disequilibria'.[29]

According to this awkward logic, a nation would increase its GDP if we paid our neighbours to look after our children and do our laundry, and they paid us to do theirs.[30] Underlying this 'common-sense' approach to household work is the utility theory of value: what is valuable is what is exchanged on the market. The implicit production boundary is determined by whether money changes hands for the service. Therefore, there is 'extreme difficulty' in giving a value to work done by women (or men) who do not receive a wage in exchange for it.

By contrast, it is remarkable how national accountants go to great lengths to include inside the production boundary the house itself, the property in which the supposedly unproductive household work is done. In the national accounts, houses owned by their occupants generate services that are included in GDP. In the US, such 'work' contributes 6 per cent of GDP – that is, a cool $1 trillion – even though none of these dollars actually exist.

How do the statisticians come up with such an absurdity? They impute a rent to everyone who lives in their own home. A market rent is estimated for a property which the owner-occupier then pays herself as lessor for the services the house provides. Since the imputed rent is regarded as income, it is also recorded in the national accounts as production. Accountants

justify this with the argument that ‘both international and inter-temporal comparisons of the production and consumption of housing services could be distorted if no imputation were made for the value of own-account housing services’.[31]

How might this work? Let’s contrast two countries. In one, there are only renters paying owners such as real-estate companies (in Switzerland in 2014 more people lived in rented homes than in owner-occupied homes). In the other, all houses are owned (in the US and UK a larger percentage of people own than rent). Since real estate adds value and income (rent) from the actual rent charged (as opposed to the ‘imputed rent’ calculated), the first country would have an unfairly high GDP compared to the other, at least in terms of the percentage of GDP deriving from property.

From a different perspective – one that sees no greater value in renting over owning a house, especially when there is no rent control – we could equally well ask why real-estate rents should add value in the first place. Another valid question is why a hike in rent should increase the value produced by real-estate agencies, especially if the quality of the rental service is not improving. London and New York City tenants, for example, know only too well that property management services do not improve even though rents rise – in London’s case, rapidly in recent years.[32]

It’s also worth noting that the national accounts treat property and real estate (both residential and commercial) as comparable to a firm. Buying a house or factory building is called an ‘investment’. It is assumed that the owner goes on ‘servicing’ the building, investing in its upkeep or improvement, so their income is ‘payment for a service’ and not just rent. Capital gains on buying and selling property are treated like those that apply to a business or a financial asset – although the extent to which a building is ‘productive’ is debatable. Capital gains from holding property arise out of increases in land value, which itself are determined by collective investment (in roads, schools, etc.) – little to do with the effort of the property owner.

As with the absurdity of neighbours paying each other to do their housework, it is as if the statisticians are saying that a nation of owner-occupiers could artificially amplify GDP by swapping homes with their neighbours and paying rent to one another. Statisticians have fiercely defended their treatment of income from property. But when real-estate prices appreciate rapidly, as in the US and the UK before 2007 and in hot-spots such as London even after the financial crisis, there are alarming implications for measuring. Rising house prices mean rising implicit

rentals, and hence rising incomes when the implicit rental is included. The paradoxical result is that a house price bubble, perhaps caused by low interest rates or relaxed lending conditions, will show up as an acceleration of GDP growth. Why? Because households' services to themselves – as their own landlords, charging themselves implicit rentals – are suddenly rising in value, and that is counted as income which adds to GDP. By the same token, if you strip out those imputed rentals, GDP can be shown to have risen more slowly in the years before the financial crash than after 2009.[33]

Prostitution, Pollution and Production

So national accountants' approach to valuation affects the production boundary, sometimes in intriguing ways. In the Netherlands, where prostitution is legal and regulated, the tax authorities have asked sex workers to declare their earnings, which count towards national income. In other countries, such as the UK, earnings from prostitution are not included in national income, except perhaps in estimates of the black economy.

Equally importantly, the boundary loops around the issue of the environment. Consider a river polluted by industrial waste. When the polluter pays to clean it up, the expenditure is treated as a cost which reduces profits and GDP. But when the government pays another company to clean up the river, the expenditure adds to GDP because paying workers adds value. If the cost of cleaning up pollution is borne by someone other than the polluter it is called an externality – the cost is 'outside' the polluter's profit-and-loss account – and increases GDP. Kuznets argued that such a calculation should be balanced by the 'disservice' that has been created by pollution, and therefore that the cost of that 'disservice' be taken out of the 'net' calculation of value added. But national accounts do not do that: instead, they state that it is not 'appropriate' or 'analytically useful' for 'economic accounts to try to correct for presumed institutional failures of this kind by attributing costs to producers that society does not choose to recognize'.[34]

National accountants present this question of whether something is 'analytically useful' or not as a vague argument, without reference to value. To be fair, they also rightly caution that it would be extremely difficult comprehensively to cost such externalities – negative or positive 'side effects' of production – which are not priced. All of which just

highlights the difficulties of being consistent and drawing a clear production boundary.

So while Marshall claimed that Nature does not make jumps (recall the discussion of *natura non facit saltus* in Chapter 2), national income, it appears, can do so! If self-employment (referred to as own-account production for small farming or sex workers, for example) grows in importance, or if a way can be found to cost externalities, national income will jump when the statisticians decide to include it.

The Black Economy Gets into the SNA

Something similar happens with the black or – to use the official euphemism – 'informal' economy when countries decide that it has grown so large that they must start to include estimates of it in national accounting. Consider Italy, a 'developed' country. The Group of 7 (G7), the international club of the biggest economies, estimates that in 2015 the informal economy made up 12.6 per cent of Italy's GDP.[35] That calculation excludes illegal activities, which Italian statisticians decided to leave out of their GDP measures. Since the Great Recession which began in 2008, many more unemployed Italians have taken up informal production. The Organization for Economic Cooperation and Development (OECD), the grouping of mainly high-income countries, estimated that in 2013 Italy's black market (including illegal activities – around 1 per cent of GDP) was a massive 21 per cent of GDP.[36] The same study found that, across other European countries, informal activities comprised between 7 per cent and 28 per cent of GDP – activities which were incorporated in the national accounts upon the recommendation of the 1993 and 2008 SNA.

All this begs the question: where does one start and stop? What is, or is not, to be included in the national accounts? The very fact that these questions are so difficult to answer illustrates the idiosyncrasies and vagaries of the accounting system. And the biggest oddity of all has turned out to be the so-called 'banking problem': how to estimate the productiveness of finance.

More than any other sector, finance highlights the arbitrary way in which modern national accounting decides where to draw the production boundary. When the financial sector was small (before its boom in the 1970s), there was little difficulty in excluding it; interest was as much a question of morality (positions against usury) as of economics. But as the size of the financial sector grew it became more awkward to exclude it

from national output. The tension between economists' – and indeed society's – long-held views of banks as unproductive and the steady post-war growth of the sector gave rise to what is known as the banking problem.

Until the 1970s, one of the principal sources of banks' profits – net interest payments, which are the difference between the interest that banks charge for loans they make and the interest they pay on deposits – was excluded from output in the national accounts. The only part of banks' income which was included was fees for services people actually paid for, such as the cost of opening or closing a bank account or getting mortgage advice.

Yet next came an extraordinary change. From being perceived as transferring existing value and 'rent' in the sense of 'unearned income', finance was transformed into a producer of new value. This seismic shift was justified by labelling commercial bank activities as 'financial intermediation', and investment bank activities as 'risk-taking'. It was a change that co-evolved with the deregulation of the sector, which also swelled its size even further. As this part of the story – how finance has been 'accounted' for – is too big to treat in this chapter, the next two will be devoted to it.

Profits versus Rents

As we saw in Chapter 2, the discussion about which parts of society were productive or unproductive was much less explicit before the arrival of marginal utility theory. And as we have seen in this chapter, moreover, as long as products and services fetch a price on the market, they are deemed worthy of being included in GDP; whether they contribute to value or extract it is ignored. The result is that the distinction between profits and rents is confused and value extraction (rent) can masquerade as value creation.

The complexity of assessing government value added pales in comparison with this glaring weakness in the SNA: a confusion between profits and rents. Disentangling the two is fundamental to understanding value. As we saw earlier, classical value theory held that income from activities outside the production boundary was unearned. Rent – which was regarded as unearned income – was classified as a transfer from the productive to the unproductive sector, and was therefore excluded from GDP. But if, as marginal utility holds, the 'services' of a landlord or hedge

fund manager are treated as productive, they magically become part of GDP.

The SNA generally links what people earn with the industries which pay them. Steel workers are paid by steel makers, shop workers are paid by retailers, insurance workers are paid by insurance companies, and so on. But income from property, dividends and lending, for example, is different because the people receiving it are not necessarily directly linked to its source (rent, dividends, loan interest etc.). If a steel firm rents an office, the rent it pays could go to firms in other sectors, to the government or even to households. A rich investor can derive income from dividends paid by any number of productive companies. A creditor – such as a bank – can lend money to several businesses or households and receive income as interest from them. All these types of income cannot be pinned down easily in the product account.

Although the SNA 2008 tried to deal with this difficulty, it did not state that, for example, property income is a reward for production, merely that 'Property income accrues when the owners of financial assets and natural resources put them at the disposal of other institutional units.'[37]

PATCHING UP THE NATIONAL ACCOUNTS ISN'T ENOUGH

So while, in theory, balancing the national accounts between income and expenditure requires a clear sense of where the production boundary lies – where value is created – in practice the boundary is far from clear. National accounts as they stand are certainly much better than nothing and, among their merits, do permit consistent comparison between countries and over time. But despite all the effort that has gone into developing it, the SNA lacks a coherent and rigorous underlying value theory.

Government agencies such as the Bureau of Economic Analysis (BEA) in the United States and the Office of National Statistics (ONS) in the United Kingdom employ armies of people to estimate GDP, making decisions about what is producing new value that enlarges the wealth of a country. We are mesmerized into seeing this as the domain of a highly specialized profession that uses sophisticated modern statistical methods to provide precise parameters for the value that our society produces. The growth rate of our economies is forecast years in advance using complicated mathematics, with potential 'outputs' measured and GDP estimated to a tenth of a percentage every quarter.

In reality, the national accounts have been subjected to repeated attempts to patch them up and make them more relevant to changing needs and economies. Accounting for environmental damage has been

mentioned. Accounting for happiness is another case. Lest the idea seems impossible, or at least nothing to do with economics, it's worth recalling something basic: there is no point to the economy unless it helps people to lead better lives – and that quite reasonably means, at least in part, happier lives. The American economist James Tobin (1918–2002), who won the Nobel Prize in Economics in 1981 and was a Professor of Economics at Yale University for many years, wrote:

> The whole purpose of the economy is the production of goods and services for consumption now or in the future. I think the burden of proof should always be on those who would produce less, rather than more, on those who would leave idle men or machines or land that could be used. It is amazing how many reasons can be found to justify such waste: fear of inflation, balance-of-payments deficits, unbalanced budgets, excessive national debt, loss of confidence in the dollar.[38]

Making decisions about which goods and services to include in GDP involves returning to the concept of the production boundary at the centre of classical economic thought – distinguishing productive from unproductive activities – and the theory of value that justifies such a distinction. Is a theory just assumed, as in the work of Petty and King? Or is it spelled out, as in Marx? And how can national accountants be persuaded that an activity that was previously seen as a transfer of existing value is actually creating new value? And above all, what do we mean by growth?

The way we define and measure growth is of course affected by our theory of value. And the resulting growth figures may guide the activities that are deemed important. And in the process possibly distort the economy.

GDP is worrying citizens and politicians everywhere: is it going up? Falling? And by how much? Understanding how GDP is constructed is thus crucial.

Unlike statisticians at the time of Smith or Marshall, modern governments have a wealth of data and a sophisticated system of national accounts that tracks the economy and the growth of each of its sectors. On the one hand, this makes it possible to see in great detail who does what in the economy – who is a 'value creator' and just how much everyone contributes to the national product. On the other hand, because of the way in which these accounts are set up, they are no more an objective metric of value than Quesnay's categorizations, or Smith's, or Marx's.

In essence, we behave as economic actors according to the vision of the world of those who devise the accounting conventions. The marginalist theory of value underlying contemporary national accounting systems leads to an indiscriminate attribution of productivity to anyone grabbing a

large income, and downplays the productivity of the less fortunate. In so doing, it justifies excessive inequalities of income and wealth and turns value extraction into value creation.

Put bluntly, any activity that can be exchanged for a price counts as adding to GDP. The accountants determine what falls into this category. But what criteria do they use? The answer is a hodge-podge which combines marginal utility with statistical feasibility and some sort of common sense that invites lobbying rather than reasoning about value. It is this that determines where the production boundary is drawn in the national accounts.

The next chapter focuses on the most egregious case of boundary-hopping: that of the financial sector, formerly seen as unproductive, now a creator of value.