

---

## The Economic Approach to Tax Design

This is not a book about how much public spending or how much redistribution there should be. Nor are we addressing the question of what is the right total level of taxation. But if we are to have public spending, we must also have taxation. And taxes are certainly not costless. It is impossible to take 40% or more of national income in tax—as most advanced economies do—and not have major economic impacts. Most taxes influence people’s behaviour in unhelpful ways and all reduce the welfare of those who bear their economic burden. The challenge for tax design is to achieve social and economic objectives while limiting these welfare-reducing side effects.

Most of this book looks at particular aspects of the tax system—how it treats earnings, savings, consumption, companies, housing—and asks how taxes can be designed to minimize their negative effects on welfare. To understand these particular effects, we need a framework for thinking about how to judge a tax system and how to think about its effects on welfare, distribution, and efficiency. These are the subjects of this chapter.

We begin by looking at how we might evaluate a tax system, in particular with respect to its impact on distribution and on economic efficiency. Starting with a clear understanding of what our objectives might be is crucial, and the issues are not altogether straightforward. We then move on to the really important insights of the economic approach to tax design and ask how we achieve and trade off different objectives. That is the focus of Section 2.2, where we introduce the optimal tax approach to tax design and also introduce some important ‘rules of thumb’, one of which—a

presumption in favour of neutrality—plays a very important role throughout this book.

## 2.1. ASSESSING TAX SYSTEMS

How do we identify a good tax system when we see one? One way is to see how it stacks up against a ‘checklist’ of desirable properties. The most famous is the four canons of taxation set out by Adam Smith in *The Wealth of Nations*:

- (i) The subjects of every state ought to contribute towards the support of the government, as nearly as possible, in proportion to their respective abilities ...
- (ii) The tax which the individual is bound to pay ought to be certain and not arbitrary ...
- (iii) Every tax ought to be levied at the time, or in the manner, in which it is most likely to be convenient for the contributor to pay it.
- (iv) Every tax ought to be so contrived as to take out of the pockets as little as possible, over and above that which it brings into the public treasury of the state.

These recommendations may command near-universal support but they are not comprehensive, and they do not help with the really difficult questions which arise when one objective is traded off against another. The way we formulate the objectives of a tax system is to say that *for a given distributional outcome*, what matters are:

- the negative effects of the tax system on welfare and economic efficiency—they should be minimized;
- administration and compliance costs—all things equal, a system that costs less to operate is preferable;
- fairness other than in the distributional sense—for example, fairness of procedure, avoidance of discrimination, and fairness with respect to legitimate expectations;
- transparency—a tax system that people can understand is preferable to one that taxes by ‘stealth’.

As we will see below, simple, neutral, and stable tax systems are more likely to achieve these outcomes than are complex, non-neutral, and frequently

changing systems. But simplicity, neutrality, and stability are desirable because they promote these ultimate outcomes, not in their own right.

A good tax system will not just limit negative effects on efficiency. It will also promote economic welfare by dealing with externalities which arise when one person or organization does not take account of the effects of their actions on others. Taxes can affect this behaviour by altering the incentives for certain sorts of behaviour, most notably when polluting activity is taxed to reduce the total amount of pollution.

We have formulated the question of the assessment of a tax system by suggesting what to take into account *given* a desired distributional outcome. So understanding how to think about the impact of the tax and benefit system on the distribution of income (or welfare) is clearly central. We look at that first. We then focus on the effects of the system on economic efficiency. This is the most important constraint on tax system design. We then turn to issues of fairness and transparency, and the other positive effects a tax system can have on correcting market failures such as externalities.

### **2.1.1. The Impact of the Tax and Benefit System on the Distribution of Income**

People differ, of course, in the extent to which they value redistribution. But assessing the degree to which redistribution is achieved by any given tax system is by no means easy.

The redistributive impact, or progressivity,<sup>1</sup> of a tax system is often judged by looking at how much tax individuals or households pay relative to their income over a relatively short time period—rarely more than a year. But people's incomes tend to change over their lives, which means that this approach can be a poor guide to how progressive the tax system is relative to a person's lifetime income. This is important in practice.

Focusing on snapshots of current income can paint a misleading picture. A tax change that hits someone who is earning a lot this year will seem progressive. But if this is an unusually good year for the person in question,

<sup>1</sup> Progressivity has a particular meaning for economists, set out in Box 2.1.

**Box 2.1.** Progressivity

There is a strict economic definition of progressivity. A tax is said to be progressive when the average tax rate rises as the tax base rises. So an income tax is progressive when the average tax rate rises as income rises. (We usually think in terms of annual income, though lifetime income may be the better base against which to assess progressivity.) This is the case when the marginal tax rate (the proportion of an additional pound of income paid in tax) is higher than the average tax rate (the proportion of total income paid in tax). In effect, the higher marginal tax rate pulls the average rate up towards it.

The simplest way to achieve progressivity in an income tax is to have a tax-free allowance before tax starts being payable. To see this, suppose the first £10,000 of income is free of tax and all further income is taxed at 20%. Someone earning £20,000 has a marginal tax rate of 20%. Their average tax rate is 10%.<sup>a</sup> Someone earning £100,000 would still face a marginal rate of 20%, but their average rate would be 18%. Thus a *flat tax*—an income tax charged at a single constant rate above a tax-free allowance—is progressive, as long as there is a tax-free allowance. This income tax can be made more progressive by (i) increasing the tax-free allowance, (ii) increasing the single rate of tax, or (iii) introducing one or more higher marginal tax rates on higher incomes. Progressivity does not, however, require that the marginal tax rate keeps on increasing as incomes rise.

<sup>a</sup> The tax payment is £2,000, which is 20% of £20,000 – £10,000. This gives the average tax rate as  $(2,000 / 20,000) \times 100 = 10\%$ .

then the lifetime effect may be quite different. Variation in earnings across years is not uncommon. Over a lifetime, earnings tend to start low when young, rise for a period of time, and then flatten off or fall in later years until retirement. However, this pattern is highly variable among people, depending on the nature of their occupation and skills. Ideally, we should judge the distributional impact of the tax system over a lifetime rather than at a point in time.

To illustrate this variation, consider a policy aimed at increasing taxes on the 1% of the population with the highest incomes. Generally, around 4% of 45- to 54-year-olds are in this group at any time, which suggests that *at least*

**Table 2.1.** Position of individuals in the income distribution by quintile in 2008 in relation to 1991

		Position in 1991				
		Bottom quintile	Second quintile	Third quintile	Fourth quintile	Top quintile
Position in 2008	Bottom quintile	34	23	18	15	10
	Second quintile	25	26	21	18	11
	Third quintile	18	22	21	20	18
	Fourth quintile	16	17	21	23	23
	Top quintile	8	12	18	25	38

Source: Department for Work and Pensions, 2010e, table 4.1 (BHC).

4% of people are in the highest-income 1% at some point in their lives.<sup>2</sup> Table 2.1 illustrates the point further by showing how people moved around the income distribution between 1991 and 2008. The population is divided into quintiles (fifths) from poorest to richest and the table shows how position in the distribution changed over the 17-year period. For example, whilst 34% of those in the poorest quintile in 1991 were also in the poorest group in 2008, 8% of them had made it into the top quintile by 2008 and a further 16% were in the second-richest group. Conversely, while 38% of those in the richest group in 1991 were still there in 2008, 10% were in the poorest quintile by then and 11% in the second-poorest.

These lifetime variations help to explain why the treatment of savings is so important when deciding how to tax—and why such attention is paid to the issue in this volume. People accumulate and run down savings and debts to smooth their spending over time and, as we shall see in Chapter 13, it is in principle possible to treat savings so as to approximate taxing people on their lifetime income.

While we might ideally like to know how people's lifetime tax payments vary according to their lifetime resources, we cannot usually observe either lifetime tax payments or lifetime resources. However, the use of current expenditure alongside current income can help in assessing the degree of redistribution achieved by the system. If people borrow and save to maintain

<sup>2</sup> Source: Our calculations based on Brewer, Saez, and Shephard (2010).

a stable level of consumption in the face of varying income, then current expenditure may be a better guide to lifetime resources than current income. Of course, consumption needs also vary over the life cycle. Often, current income and current expenditure provide complementary indicators; much can be gained by looking at both and considering what each measure reveals. As well as the use of income or expenditure to assess whether a household is 'rich' or 'poor', we must also consider whether to assess the household's tax payments as a proportion of its income or its expenditure. It can be misleading to look at current payments of all taxes as a percentage of current income: in general, a better guide to the lifetime distributional impact is to look at income taxes as a percentage of current income and expenditure taxes as a percentage of current expenditure. In the absence of data on whole lifetimes, snapshots of current income and expenditure must be used judiciously to give a rounded impression of the distributional impact of taxes.

In assessing progressivity, we should also look at the impact of the system as a whole rather than at its individual components. This assessment should include welfare benefits and tax credits because it is the overall effect of the tax and benefit system that matters. Making the system as a whole progressive does not require every individual tax to be progressive. Different taxes are designed to achieve different ends. Some current taxes are quite regressive—taxes on tobacco, for example—because they are intended to achieve a different purpose, not progressivity. This is an issue we return to in Chapters 6 and 9, where we consider the structure of VAT and argue that zero-rating goods consumed disproportionately by poorer households is not a good way to achieve progressivity in the tax system as a whole.

A further issue concerns the need to think about whether to judge the redistributive effect (and other effects) of taxes on an individual or a household basis. Economics has developed a comprehensive theory of individual behaviour. But most people live in households—with spouses, partners, children, or parents. Over time, the structure of households is fluid because of birth, death, and break-up. This poses difficulties for tax design. Current UK practice is not consistent: income tax and National Insurance contributions are levied on individuals, while benefits and tax credits are paid to households. Economic theory has similar difficulties in reaching a

consensus on the right approach.<sup>3</sup> The issues that are raised by the individual/household distinction cannot be easily resolved. We consider them a little more in the context of direct taxes and benefits in Chapters 3 and 5.

### *The Burden of Taxation*

An even more fundamental question regarding the redistributive effect of taxes relates to how we assess who is actually bearing the economic burden of the tax. That need not be the person or organization that makes payment to the tax authority. Nor need it fall on the *statutory bearer*—the person or organization legally responsible for the tax. To illustrate the difference between these two concepts, note that under the Pay-As-You-Earn (PAYE) system in the UK, the worker is the statutory bearer of income tax, but the firm remits the tax. The allocation of the statutory burden of taxation among taxpayers is called the *legal incidence*. The legal incidence of a tax can be very different from the *economic incidence* or who *bears the burden*. An individual bears the burden of a tax to the extent it makes him or her worse off (that is, causes a loss of welfare).

The fact that income tax is remitted by employers does not alter the fact that at least part of the economic incidence falls upon employees. Similarly, although employers and employees formally pay separate National Insurance contributions on the employees' earnings, the eventual economic incidence of the contributions is likely to be the same. Employers make decisions based on the total cost of each employee, and employees are interested in their salary after tax. In the long run at least, the allocation of National Insurance contributions between employee and employer should make no difference to the number of people the employer chooses to employ or the after-tax wage of the employee.

In many cases, the economic incidence of a tax can be hard to identify. For example, when the excise duty on alcohol is increased, the price of drinks in the shops need not rise by the amount of the tax increase. Firms that produce and supply alcohol may choose to absorb part of the tax increase and only pass a fraction of it on to consumers. Or consider stamp duty on the sale of a house. It is natural to assume that, since stamp duty is paid by the purchaser,

<sup>3</sup> The literature is surveyed by Apps and Rees (2009).

it does not make the seller any worse off. But this is wrong. Suppose there is a completely fixed supply of houses. The price of a house will then be determined by demand—the willingness of buyers to pay a certain amount. Imposing stamp duty will not change the total (price plus stamp duty) that buyers are willing to pay, so the price must fall: the economic incidence falls on the seller of the property, not the purchaser.

Allocating the economic incidence between firms and consumers is only the first step of the analysis. The second step follows from observing that firms are just legal entities and enjoy no economic well-being beyond that of their customers, employees, and shareholders. The impact of any tax paid by a business—either in the sense that the tax is remitted by the business or in the sense that it is the statutory bearer of the tax—can be traced ultimately to a reduction in the economic welfare of the owners of the business, the suppliers of capital and other inputs to the business, and/or the employees. So, even though the first step allocates part of the incidence to firms, the burden must ultimately be borne by some combination of customers, employees, and shareholders. The final distribution of the burden is nearly always unclear to the individuals concerned, and often difficult for economists to determine. This no doubt helps explain why, in virtually all countries, the statutory liability of employer social insurance contributions is much greater than that of employee contributions, and why taxes on firms are often seen as ‘victimless’. They are not. They are in the end paid by customers, employees, and shareholders. We cannot emphasize this point too much.

### **2.1.2. The Effect of Taxes on Economic Output and Efficiency**

Prices play a central role in the modern market economy, as signals that reflect and guide the decisions of consumers and producers. Taxes disrupt these signals by driving a wedge between the price paid by the buyer and the price received by the seller. For example, income tax means that an employer pays more for an hour of work than the employee receives for it, while VAT means that a retailer receives less for a product she sells than her customer pays for it.

By increasing prices and reducing quantities bought and sold, taxes impose losses on consumers and producers alike. The sum of these costs almost always exceeds the revenue that the taxes raise—and the extent to which they do so is the *deadweight loss* or social cost of the tax. A key goal for tax design is to reduce the deadweight loss of the system as a whole as far as possible.

The size of the deadweight loss is related to the *elasticities* of demand and supply for the item being taxed—in other words, the extent to which demand and supply respond to changes in price.<sup>4</sup> The more elastic is the demand for a product with respect to its price, the more a given tax increase will reduce demand for it. High elasticities therefore mean large deadweight losses for a given change in tax.

There are generally two channels through which tax changes influence behaviour: the income effect and the substitution effect. Consider the impact of an increase in a tax on earnings on people's work decisions. The tax will reduce the income that people receive for a given number of hours of work, encouraging them to work more to limit any decline in living standards. This is the *income effect*. But the tax will also make an hour of work less attractive relative to an hour of leisure than it had been previously, encouraging people to work less. This is the *substitution effect*. The two effects work in opposite directions. For any individual tax change, it is impossible to say a priori which will dominate. But for revenue-neutral changes to taxes, income effects will tend to roughly balance out on average: money given away to some people must be matched by money taken from others, so a positive income effect for one group will be offset by a negative income effect for another (though it is possible that one group might be more responsive than the other). Substitution effects will not necessarily balance out in this way.

In addition, most empirical work<sup>5</sup> suggests that it is the substitution effect that dominates. So additional taxes on labour earnings typically reduce hours of work. But the strength of these effects differs between different types of worker. Often they affect behaviour less at the *intensive margin*—whether to

<sup>4</sup> If an increase in price prompts no change in behaviour, then the elasticity of that behaviour with respect to price is 0. If a price change leads to a behavioural change of the same magnitude—e.g. a 10% increase in the price leads to a 10% fall in demand—then the elasticity has value  $-1$ . An elasticity of  $-0.1$  means that a 10% increase in price reduces demand by only 1%.

<sup>5</sup> See e.g. Blundell and MaCurdy (1999) and Meghir and Phillips (2010).

work slightly more or less—than at the *extensive margin*—whether to undertake paid work at all. This is an important insight which we dwell on in the next chapter.

It is not just income taxes that have income and substitution effects on decisions about working. The introduction of a uniform VAT means that less can be purchased with any given income (giving rise to an income effect) and that each hour of work can purchase fewer goods (leading to a substitution effect). What matters in deciding whether and how long to work is what can be purchased with the earnings received. So increasing taxes on goods and services reduces the pay-off to working more in the same way that increasing direct taxes does—a fact sometimes overlooked by those who argue we should move away from income tax and towards indirect taxes to preserve work incentives.

Measuring these effects is crucial if we are to assess the impact of tax changes. In general, we would want to levy taxes at higher rates where they will have less effect on behaviour. If there are groups of people who respond to high tax rates more than others, then we might want to shape the tax system accordingly. A ‘one size fits all’ tax system is economically costly when people respond differently to tax changes. The question is whether the benefit of designing the tax system in a more tailored way outweighs the operational and compliance costs of a more complicated system. One needs to be sure that all the possible behavioural effects are well measured and then trade off the gains from differentiation against the costs. In many circumstances, there is sense in demanding a high standard of proof that divergences from uniform treatment are worthwhile.

Importantly, taxes can have long-term or *dynamic* effects as well as immediate effects on behaviour. We know, for example, that higher petrol taxes have more effect on petrol consumption and less effect on miles driven in the long run than they do in the short run, because people respond by demanding more fuel-efficient cars and manufacturers respond by supplying them. We know rather less about the dynamic effects of some other important policies. For example, benefits supporting low earners, such as the Working Tax Credit in the UK, are effective at getting more people into

work in the short run. But what effect do they have in the long run, perhaps by influencing the decisions people make about education or training? We know little about these effects and they may be more significant than many of those that we do understand much better.

That there are major effects of taxes on behaviour is undeniable. This fact has led economists to think about the potential for two sorts of taxes that should not have these effects—lump-sum taxes and taxes on economic rents:

- Lump-sum taxes are taxes liability to which cannot be altered by any change in behaviour. In principle, the liability to a lump-sum tax can be linked to any unalterable individual characteristic (such as a taxpayer's age or some measure of inherent ability). In practice, such characteristics tend to be either undesirable as the sole basis for determining someone's liability or impossible to define or measure. Levying the same lump-sum tax on everyone is, in principle, feasible, but the lack of any link between the tax and ability to pay renders it politically unattractive, to say the least (even though we should care about the redistributive consequences of the tax system as a whole rather than of any individual component of it). Public reaction to the UK community charge, to give the poll tax its official title, is a graphic illustration of this point.
- Taxing pure economic rents does not create a distortion. An economic rent arises when a resource generates a high return relative to its next-best use. When a rent arises, taxing it should not alter behaviour, since only the excess income over the next-best use is taxed. Rent is most often associated with the return to land. Land derives its value from its location and this makes the return to owning land attractive to tax, because the owner cannot move it elsewhere. We address the issue of the taxation of land value in Chapter 16. Rents may also arise from the right to extract scarce and valuable natural resources and among innovators, artists, sports stars, and firms with recognizable brands.

Neither lump-sum taxation nor the taxation of rents is a terribly helpful guide to most policymaking. The fact is that most taxes will alter behaviour and reduce both welfare and economic output.

### **2.1.3. Taxes Designed to Change Behaviour**

In general, taxes reduce welfare. But there are situations in which taxes are deliberately used to change behaviour in order to promote welfare. Environmental taxes designed to tackle spillovers, or externalities, are the most obvious examples. Where one economic agent imposes costs (or indeed benefits) on others and does not take account of those effects when deciding on his actions, the tax system can be used to ensure that he ‘internalizes’ the costs he imposes. In other words, one can use a tax to provide a price signal which would otherwise be absent. It is because the tax can stand directly in place of a market price signal that taxation can be a particularly efficient way of dealing with externalities. We discuss these issues in some detail in Chapter 10 and with specific reference to climate change and to motoring in Chapters 11 and 12 respectively.

It is not just with respect to reducing environmental damage that taxes are used to alter behaviour. Excessive drinking can cause costs to society, and most countries impose taxes on alcohol well above the usual rates of consumption tax. High taxes are also imposed on tobacco. In fact, particularly in the case of tobacco, these taxes are not wholly, or perhaps even largely, designed to reduce harm to others. They are imposed in recognition of the harm that individuals can do to themselves, particularly in the face of addictive substances.

There is also likely to be a large dose of straightforward paternalism in such taxes. Governments (rightly or wrongly) do take a view as to how they would like their citizens to behave and adjust tax systems accordingly. This may be the best explanation for the zero-rating of books for VAT purposes in the UK.

Some actions by individuals or companies can also create benefits for the wider economy that are not recognized in the price signals they face. Again the tax system can help to compensate for the absence of such price signals. A good example is the R&D tax credit, which is found in a number of countries and provides a subsidy to companies investing in research and development (R&D). Any one company’s R&D is likely to create positive spillovers for the rest of the economy. But companies won’t take that into account in making their own decisions. A tax incentive might give them the price signal to do an amount of R&D closer to that which they would have

done if they were taking account of the positive spillover effects of their own decisions. Tax systems that recognize these sorts of spillovers effectively may have quite significant effects on growth and welfare. As we suggest in Chapter 12, for example, there are potentially large environmental benefits associated with a tax system better designed to capture the externalities associated with driving.

#### **2.1.4. A Fair and Transparent Tax System**

Whether the tax system is seen as 'fair' is not simply a question of redistribution. Fairness of procedure, fairness with respect to legitimate expectations, and fairness in treating similar people similarly also matter:

- A tax system is more likely to command respect, and so be widely accepted, if the process that determines tax levels and structures is seen to be fair. This is what we mean by fairness of procedure. The process and institutional context for tax policy matter, not just because they are likely to determine the outcome, but also because they affect how that outcome is perceived and even how well it is complied with. A process of policy determination is needed that ensures that even those who lose out accept the legitimacy of the outcome. Workable democratic procedures have this at their heart and are supported by a process of debate and consultation. We do not spend much time on this issue in this volume, but its importance should not be underestimated.
- Another sort of fairness, which can be related to this concept, is that of fairness with respect to legitimate expectations. Tax changes that impose unexpected losses relative to previous expectations can be perceived as 'unfair'. This is most often true of capital taxes, which, for example, might reduce the net expected capital gain from an asset, or indeed might reduce the value of assets into whose value a particular expectation of the tax rate has been capitalized. In fact, any tax change can have this kind of effect—my investment in my own human capital may become less valuable as a result of an increase in income tax, in just the same way as my house may become less valuable as the result of a tax change. While legitimate expectations matter, effects of this kind can be very hard to avoid and have to be weighed in the balance against potential longer-term benefits.

- More closely related to the idea of distributional fairness is the notion that the tax system should treat similar people in similar ways. This has been dubbed *horizontal equity* by economists. The difficulty with applying this concept in practice is defining similar individuals, and different countries define them in different ways. For example, the UK taxes individuals with similar earnings in similar ways whether or not they are part of a married couple. In contrast, the US taxes married couples with similar earnings in similar ways regardless of how individual earnings are split between the partners.

It is not even obvious that the tax system should treat people with similar levels of earnings in similar ways. If one person can earn £500 a week in 20 hours, then she is better off than someone who needs to work 40 hours to earn the same amount. Perhaps she should pay more tax. People also differ in their needs—perhaps because of health status or number of dependent children—and tax systems might reasonably differentiate accordingly. In France, for example, the number of children influences the amount of tax paid.

Ideas of fairness can also be applied to the range of economic activities that are taxed. A neutral tax system—one that taxes similar activities similarly—avoids giving people encouragement to shift from high- to low-taxed activities in a way that is economically costly. It also avoids discriminating between people who make different (but inconsequential) choices. There is unlikely to be any legitimate case for taxing silver cars at a higher rate than blue cars—that seems simply unfair. There are aspects of different VAT rates in the UK tax system which can also be seen as unfair. Those who like biscuits (subject to VAT) are in this sense unfairly treated relative to those who prefer cakes (not subject to VAT). Neutrality between goods can promote fairness as well as efficiency.

None of these senses of fairness is absolute. More or less any tax proposal will face some charge of unfairness viewed from some perspective. But this is best dealt with by trying to be open and transparent about the arguments and evidence that underpin the proposal. Tempting though it may be to disguise who gains and who loses, in the long run the cause of sensible reform is best served by being honest about the objectives and consequences of what is being proposed. This is among the reasons for seeing transparency as an important part of a good tax system. Lack of transparency can easily

lead to poor process, to outcomes that lack fairness in some dimension, and eventually to a lack of legitimacy, which can fuel non-compliance.

## 2.2. ACHIEVING AND TRADING OFF OBJECTIVES

We want a tax system that does not unnecessarily discourage economic activity, that achieves distributional objectives, and that is fair, transparent, and administratively straightforward. How can we achieve these outcomes and how should we trade them off? In particular, how should we think about balancing efficiency loss against equity?

The question of how to trade off these objectives is the subject of optimal tax theory, to which we come in the next subsection. We then look at three ‘rules of thumb’—neutrality, simplicity, and stability—which might help to guide the design of a tax system. We distinguish these from the basic criteria for a good system because, whilst generally likely to be desirable, these are not ends in themselves in the way that progressivity and efficiency are goals of the system. Rather, they are instrumental guidelines, the following of which is likely to help achieve the final goals.

Before delving into these issues, though, it is worth stressing one point which will be crucial to the messages in much of this volume—that in achieving the overall objectives of the tax system, it is important to consider all taxes (and transfer payments) together as a system. It is the redistributive impact of the system as a whole which needs to be measured and judged. Not every tax needs to be progressive.

### 2.2.1. Optimal Taxation and Social Welfare

Economists have expended much effort in the study of ‘optimal taxation’. Much of this work is abstract and mathematical, and this volume is not the place for a detailed exposition of optimal tax theory.<sup>6</sup> Nevertheless, it is important for the approach taken here. It provides a methodology for designing tax systems to achieve the best outcome given the constraints

<sup>6</sup> For which see e.g. Tuomala (1990).

faced by the government. As such, it is the foundation of the idea of 'tax by design'.

Optimal tax theory is all about the choice of a system of taxation that balances efficiency losses against the government's desire for redistribution and the need to raise revenue. It provides a way of thinking rigorously about these trade-offs, and ensuring that value judgements reflecting concerns about income distribution and well-being are made explicit while the efficiency costs of achieving that redistribution are properly taken into account. The seminal theoretical work in this area<sup>7</sup> was carried out some years ago—in the early 1970s—but its impact on practical tax design continues to build, along with important developments in both theory and applications. The theory of optimal taxation begins by clarifying the *objectives* of policy and identifying the *constraints* under which it operates. The tax system that best achieves the objectives whilst satisfying the constraints is identified as the optimum.

What matters for these value judgements is the effect of the tax system on welfare rather than only on income. With a revenue-neutral reform, our judgement of the effects will depend upon the weights we give to the welfare of people at different points in the income distribution. The resulting changes in tax liabilities will redistribute income from some points in the distribution to others and this will be welfare enhancing only if we weight welfare gains to the recipients more strongly than we weight welfare losses from the losers. If we care more about the welfare of the poor than the rich, then, other things equal, we will prefer a world in which the rich pay more tax. How much more will depend on how much we care about inequality.

We also need to evaluate behavioural changes induced by changes in tax rates. In general, welfare is lost by taxing someone who responds to that tax by reducing their work effort or by putting effort into reorganizing their affairs to reduce tax payments. These behavioural responses constrain what can be achieved via the tax system. There are costs associated with work, but if individuals would have chosen to work in the absence of taxation, then the benefits of work must outweigh those costs, and reduction in work effort as a result of taxation is welfare reducing. This loss of welfare is referred to as the *deadweight loss* or *excess burden of taxation*. It too has distributional aspects

<sup>7</sup> Mirrlees, 1971.

since the magnitude of the deadweight loss will differ across the income distribution. We can assess the magnitude of the deadweight loss at any point by considering how much more tax we could have raised from individuals at that point without leaving them worse off if, hypothetically, we had had the information needed to raise those taxes without basing them on income. The size of that loss will depend both on those individuals' wages and on the responsiveness of their behaviour to tax changes. A beneficial tax reform is one that has a positive impact on social welfare taking all of these effects into consideration.

It is worth being a little more concrete about how optimal tax theory helps to inform views about the best system.

Consider the structure of tax rates on labour earnings. A good deal of the concern about progressivity in the tax system is motivated by the fact that inequality is primarily determined by disparities in employment opportunities. A progressive system will set taxes on earnings at higher levels for higher earners. Such higher tax rates impose distortions and disincentives. But these need to be balanced against the gain in achieving progressivity. Just how far tax rates can be raised depends on how responsive earnings are to tax changes. And, as we have already seen, we may want to vary tax rates where some groups are known to be more or less responsive to taxation over the life cycle. An income tax system is optimal when the gain through desired redistribution, and raising revenue, is offset in an optimal way against the cost induced by lowering labour earnings.

Optimal taxation also helps in thinking through the right structure of tax rates on goods and services. The principle is the same as for labour earnings. There are distortions induced by taxes but there are also distributional effects. But optimal tax theory does not always support taking the latter into account by differentiating rates on commodities. And this is a perfect example of why the systemic view is important. If taxes on earnings are well designed, then they can do the heavy lifting as far as achieving progressivity is concerned.

We need to approach the taxation of savings with a similar mindset. Higher earners tend also to be higher savers. But this does not automatically imply that they should face a higher tax rate on their savings. That again depends on looking at the system and assessing whether progressivity can be achieved in a more efficient way by adjusting the rate schedule. In general,

taxing savings is an inefficient way to redistribute. As we shall see in Chapter 13, though, there may be circumstances—for example, where saving behaviour reveals something additional about the underlying earning potential of individuals—in which taxing the return to savings does become optimal.

In general, there is little clear rationale for trying to use the tax system to influence when people consume in their lifetime rather than focusing on progressivity based on the overall resources that they have available. This logic is particularly compelling when inequality is mostly generated by labour market opportunities and an optimal income tax can be designed. Particularly where inherited wealth makes a big difference to the welfare of some groups, on the other hand, taxes on wealth transfers may be part of an optimal system—an issue we explore further in Chapter 15.

A broad concept of optimality should include other considerations too. Administrative and compliance costs should ideally be brought into optimal tax calculations.<sup>8</sup> Social welfare can embody value judgements other than those associated with the distribution of welfare. Most governments put a high rate of tax on tobacco products to discourage smoking even if most smokers belong to low-income groups. Such multiple objectives imply the need to think about how a tax system performs with respect to several criteria simultaneously.

Ideally, we would begin tax design by coherently laying these objectives out and constructing the tax system that best reflects these goals.

As we have seen, the optimal tax approach emphasizes all the constraints under which the government must operate, particularly those imposed by the behavioural responses of individuals and companies to the taxes that are levied. Constraints are also imposed on the tax system by the government's *limited information*. Limited information impacts directly upon the choice of *what* to tax. For example, governments can observe people's actual earnings but cannot know about each individual's underlying ability to earn income. This is a constraint, since it would arguably be better to tax on the basis of ability to earn rather than actual earnings since the former is not subject to manipulation in the face of tax changes. On the whole, we would expect high-ability individuals to earn more. But they could choose to earn less if

<sup>8</sup> Shaw, Slemrod, and Whiting (2010) provide a very insightful analysis.

the taxes that they face are too high. Optimal taxes on earnings have to take this into account and it limits the degree of progressivity that can be achieved.

Much recent progress in tax analysis has come from a better understanding of these responses and the constraints they place on policy. In some areas, better economic understanding and better policy have gone hand in hand. But there are some constraints and behavioural responses that governments still appear to struggle to comprehend. In subsequent chapters, we shall see how the tax and benefit system has led to much lower employment rates among older men in France than in the UK (Chapter 3), how the pattern of savings is influenced by the tax treatment of different forms of savings (Chapter 13), how generous tax treatment led many self-employed people in the UK to incorporate (Chapter 19), and how the tax credit system has quite different effects on the work incentives of different groups of people (Chapter 4). There are many, many more examples of apparently unintended effects of tax design on people's behaviour.

Optimal tax theory has its limitations. It cannot readily take account of many of the concepts of fairness that we have discussed. And whilst it can be adapted to incorporate issues of administrative costs, in practice it rarely is. It is nevertheless a powerful tool and, throughout this volume, the conclusions of optimal tax theory will inform the way we discuss policy. We will see in Chapter 3 what optimal tax theory has to say about the structure of marginal income tax rates. It informs our discussion on the structure of indirect taxes generally in Chapter 6 and of VAT specifically in Chapter 9. Optimal tax theory will also be important to our discussion of the tax treatment of savings in Chapters 13 and 14 and of the taxation of company profits in Chapters 17, 18, and 19.

### **2.2.2. Rules of Thumb**

Beyond the overall objectives of a tax system, and the details of the trade-offs that optimal tax theory forces us to think about, there are several instrumental guidelines, or rules of thumb, which can help in system design. Other things being equal, a tax system is likely to be better if it is simple, neutral, and stable. But none is necessarily the right thing to aim for at all

times, which is why they are guidelines against which to assess a tax system rather than always objectives.

### *Neutrality*

A neutral tax system is one that treats similar activities in similar ways. For example, a system that taxes all consumption the same would achieve neutrality over choices that people make about what to consume. A system that treats all income the same achieves neutrality over the form in which income is received. A system that taxes all forms of savings the same achieves neutrality over the form in which savings occur. A system that imposes the same present value of tax on consumption now and consumption in the future will be neutral with respect to the decision over whether to save or consume out of current income.

So a neutral system minimizes distortions over people's choices and behaviour. In general, it therefore minimizes welfare loss. In a non-neutral tax system, people and firms have an incentive to devote socially wasteful effort to reducing their tax payments by changing the form or substance of their activities.

But the promotion of neutrality in the tax system is not always an appropriate end for policy. There are times when a lack of neutrality is valuable. We do not want to be neutral towards environmental bads—we want to tax them more than other things. There is likely to be a case for offering tax relief for corporate research and development activity. There are strong and respectable arguments for treating some forms of consumption—particularly those, such as childcare, that are complementary to work—more leniently than others. The same is true of some forms of savings. We do not, in the end, believe that pension savings should be taxed in exactly the same way as other savings.

But the tax system in the UK, like that of most modern economies, is full of non-neutralities which are hard to justify, wasteful, and ripe for reform. It distorts choices between debt and equity finance, between capital gains and other forms of capital income, between owner-occupied housing and other assets, between different forms of remuneration for work effort, between different forms of carbon emissions, and between different forms of business

organization. These distortions create complexity, encourage avoidance, and add costs for both taxpayers and governments.

They are not generally inevitable and often stem from the lack of clear underlying economic principles in the design of the tax system. The problems that stem from differential treatment of debt and equity under the corporate and personal tax systems illustrate this point, reflecting the lack of any economic principle that distinguishes these two forms of business finance. Similarly, if the tax system draws unnecessary dividing lines between items that are difficult to differentiate or are close substitutes—such as dividends and salary for owner-managers of small firms or cash and fringe benefits for National Insurance contributions—taxpayers will be gifted opportunities to select the more favourable tax option. In these circumstances, governments commonly respond with an over-elaborate, tangled web of legislation that seeks precise definitions to minimize avoidance opportunities and ends up concealing within its length and complexity whatever policy objective the legislation was originally designed to achieve.

Very often, greater neutrality leads to both greater simplicity and greater fairness. Achieving it requires a holistic view of the tax system which recognizes the interdependencies between different parts of the tax system—personal and corporate taxes, taxation of dividends and earnings, taxation of debt and equity. In particular, it requires a consistent understanding of what it is we are trying to tax—the tax base. In the UK, as in most countries, the tax base for individuals remains a mix between an income base and an expenditure base, creating distortions in people's decisions over savings. The corporate tax base both creates distortions over how firms raise funds—between debt and equity finance—and how, where, and to what extent they invest, and fails to be well aligned with the personal tax base, creating distortions over the legal form that small businesses in particular may choose.

So while there will be occasions on which we might want to diverge from neutrality, and occasionally we will argue that specific non-neutralities are justified, in general these do not align with the non-neutralities we observe in the tax system. But aiming to increase the efficiency of the tax system by moving towards greater neutrality remains a good principle in guiding reform.

*Simplicity*

It is often said that a good tax system should be a simple one. And surely a simple one is better than a more complex one if it achieves the same ends. A simple tax system is likely to be relatively transparent and impose low administrative costs.

But the world is complex enough that no tax system is likely to be truly simple. And just as there are occasions when we might want to deviate from neutrality, so there will be times when we have to accept more complexity. Indeed, the two concepts are linked—a neutral tax system will tend to be a simple one and vice versa. The less differentiation there is between the taxation of similar activities, the more neutral and the simpler the system will be.

Lack of simplicity and neutrality invites tax avoidance. Individuals and firms strategically arrange their affairs in response to changes in taxes, so as to reduce their liability even though the underlying economic nature of the activities does not change. But if complexity creates opportunities for avoidance, so avoidance activity invites a complex set of rules to close loopholes. And so the process of avoidance becomes a game of cat and mouse played between the revenue authority and taxpayers. Each revision to the system is followed by the introduction of new avoidance schemes. Some schemes are demonstrated to be illegal while others are rendered redundant by new revisions to the legislation. The process then begins anew. The perpetual revision of tax law and litigation against avoidance schemes add to complexity and to the cost of implementation. Compliance costs are increased by the search for avoidance schemes and the consequences of litigation. Avoidance activity results in an addition to deadweight loss. A tax system that minimizes avoidance activities—which a simpler and more neutral system will often, though not always, do—will tend to reduce the total economic cost of taxation.

Avoidance (and evasion—illegal activity that leads to lower tax payments) also cost the revenue authorities money directly. How much is inevitably hard to judge, but HMRC estimated the *tax gap* between actual revenues and the amount that the authorities think they should be receiving for 2008–09 at approximately £42 billion, which is 8.6% of the tax that should have been

collected.<sup>9</sup> This is similar in value to the US estimate of a direct tax gap of 14% in 2001 and a Swedish estimate of 8% in 2000. Much larger tax gaps have been estimated for developing countries.<sup>10</sup>

As well as creating direct compliance costs and opportunities for avoidance, complexity can lead to costly behavioural change. As behavioural economists have been stressing for some time, complexity can also create additional unintended consequences. People are inclined to focus on problems that are easier to solve, on today rather than tomorrow, and on what other people are doing. This is an important issue for the political economy of tax reform and is one reason for the inordinate public and political focus on one particular aspect of the system—the basic rate of income tax. Complexity makes useful and informed debate difficult. Unfortunately, governments often increase complexity as they add provisions and special cases to the tax system.

If policymakers have multiple objectives for the tax system, then a substantial degree of complexity is unavoidable. Even so, we do need to justify with quite strong evidence any move towards greater complexity. In all that follows, we start from a presumption in favour of simplicity if it is clear that doing so generates significantly greater benefits than costs.

In any case, any tax system will involve a compromise between what policymakers would like to do and what they are able to do with the information and administrative tools available. The ability to levy a tax relies on being able to measure the relevant tax base—the quantity of income, expenditure, and so on—against which tax liability is assessed. The standards currently achieved have required major investment in administrative and compliance capacity. The costs of administering the tax system, and complying with it, matter a great deal. They impose significant limitations on tax design. At one extreme, we can't tax people directly on what we can't observe directly—for example, their ability. More pragmatically, not even all forms of income are easily assessed. The incomes of the self-employed are difficult to measure, as are some forms of capital gains.

The main bases for tax have changed over time as society has evolved. In the 19<sup>th</sup> century, most revenue was raised from excise and customs duties. Taxes on income only came to provide the majority of revenues in the

<sup>9</sup> HM Revenue and Customs, 2010d.

<sup>10</sup> Schneider and Enste, 2000.

second half of the 20<sup>th</sup> century. In many developing countries, levying an income tax is very challenging and, by western standards, high proportions of revenue come from taxing what is easily taxed—typically, internationally traded products as they pass through ports. Such taxes distort trade and are particularly damaging to economic efficiency.

In the longer term, what is considered feasible may change again. Might our reliance on taxing income and spending decline once other options become available? Might changes in our capacity for making genetic assessments of ability (or longevity or willingness to work) provide a new base for taxation? Perhaps not. But it is worth bearing in mind that relying on taxes levied on income seemed equally unlikely 300 years ago.

### *Stability*

Tax systems that continually change impose greater compliance costs on those who are taxed. They lead to difficulties in making long-term plans. Lack of stability can impact negatively on investment decisions by firms and on saving and investment decisions by individuals. Changing capital taxes in particular can lead to a sense of unfairness if the current structure and rates are capitalized into asset values. For all these reasons, a stable tax system is, other things equal, preferable to an unstable one.

But this must not be an excuse for permanent inaction. There are costs associated with change. But there are also costs associated with keeping a poorly designed system in place. This is a book about tax reform and we are not writing it to conclude that the virtue of keeping everything as it is outweighs the virtue of seeking something better. As we will see later on, the scale of the welfare gains available as a result of reform is, in some cases, very large indeed.

But there is virtue in having a clear and transparent method of making changes to the tax system, and a clear long-term strategy for change. Certainty is valuable, and does not require stagnation. So the process of tax reform matters as well as the content. And the failure of successive governments in the UK and elsewhere to be clear about the long-term strategy and direction for tax policy has been very costly.

### 2.3. CONCLUSIONS

The concepts discussed in this chapter are at the heart of the economic approach to tax design and are used frequently in this volume. In the UK, many of the problems involving lost revenues, unintended consequences, and policy U-turns that have characterized tax policy in recent decades have resulted from ignoring or giving too little weight to these ideas. And the UK is not alone.

If there are three ideas worth holding on to from all those discussed in this chapter, they are:

- *The need to think of the tax system as just that—a system.* The way that different taxes fit together matters, as does being clear about the role of each tax within the system.
- *The central role of redistribution in the tax and benefit system.* The extent of that redistribution will be determined by society's preferences and the impact of the system on efficiency. The trade-off between redistribution and efficiency is at the centre of very many debates about tax policy.
- *The importance of neutrality as a benchmark.* While we don't always want neutrality, it is often valuable and will always be an important benchmark for assessing the system.

In some ways, it is remarkable that governments in rich countries manage to raise such substantial tax revenues from a largely compliant population. If they want to continue to do so, they will need to take account of these principles and ensure that the tax system is seen to be efficient and fair as part of the bargain between citizens and government. As we will see in the chapters that follow, there are real opportunities to improve the current system—to make it more coherent and efficient and often, as a result, more equitable. In some cases, this involves getting rid of obvious anomalies; in some cases, it means fundamentally rethinking the tax base; in some cases, it means taking proper account of the system as a whole; and in some cases, it just means making better use of our understanding of different groups' responses to incentives in designing the system.