

Implementation of VAT

In the previous chapter, we discussed the principles of indirect taxation. In this chapter, we focus on some of the practical issues in its implementation, and in particular in the design and operation of VAT. These issues of implementation are important to policy design, and in particular to considering possible reforms to VAT. Once we have been through the important design issues in this chapter, we will look at two sets of reforms in the next two chapters—first the specific issue of VAT and financial services, and then extending VAT coverage to other goods and services in general.

In terms of revenue raised, VAT is by a long distance the most important indirect tax in the UK, and in most other OECD countries. Since its introduction in France in 1954, it has proved an exceptionally successful form of taxation and has been adopted by many countries worldwide, including all OECD countries other than the US. Bird (2010, 363) calls it ‘unquestionably the most successful fiscal innovation of the last half-century ... perhaps the most economically efficient way in which countries can raise significant tax revenues’.

In this chapter, we first explain how VAT works and why it has such appeal. We then go on to discuss its more problematic aspects: rate differentiation and exemptions, the scope for non-compliance, and difficulties applying VAT in the context of international trade.

7.1. HOW VAT WORKS

VAT taxes all sales, whether wholesale or retail, but allows registered traders to deduct the tax charged on their inputs. It is therefore a tax on the *value added* at each stage of the production process. Since the value of the final product is the total of the value added at each stage of production, the tax base—total value added—equals the value of final sales. Consequently, the tax is in effect imposed on the value of the final product but is collected in small chunks from each link in the supply chain. VAT charged on sales to registered traders who sell on an item or use it in production can be reclaimed by the purchaser; only VAT on retail sales cannot be reclaimed. VAT therefore taxes only final consumption and leaves production decisions undistorted.

It is worth illustrating how the system works with a very simple example. Suppose firm A makes a sale to firm B for £100 plus 20% VAT—£120 in total—remitting the VAT to the tax authorities.¹ Firm B uses what it has bought to make products worth £300; £60 VAT is due when these products are sold to firm C, but B can also reclaim the £20 VAT charged on its inputs. And, similarly, C can in turn reclaim the £60 VAT on its input purchases. Firm C, a retailer, sells its products to final consumers—households—for £500 plus £100 VAT.

Table 7.1 illustrates the VAT payments in this simple supply chain from two perspectives which bring out the two key features of VAT. The top panel shows the VAT payments associated with each transaction. When firm A makes a £100 sale to firm B, A charges £20 VAT on the sale, but B can reclaim the same £20, so there is no net revenue raised from the transaction. Similarly, the £60 VAT due on B's sale to C is reclaimed by C. Only sales to final consumers generate a net VAT liability, and the total revenue raised is 20% of the value of this final consumption. No net tax is levied on intermediate inputs; the pattern of activity in the supply chain generating the final product is irrelevant to the tax burden, and so is not distorted.

¹ VAT is usually expressed in *tax-exclusive* terms: a 20% VAT rate means that liability is 20% of the price excluding VAT (20% of £100 = £20). This is unlike income tax, for example, which is expressed in *tax-inclusive* terms: as a percentage of income including (i.e. without deducting) the tax itself. A 20% tax-exclusive rate is equivalent to a 16.7% tax-inclusive rate (16.7% of £120 = £20).

Table 7.1. A simple supply chain with 20% VAT

	VAT charged on sales	VAT reclaimed on input purchases	Net VAT liability
<i>Analysis of transactions</i>			
Sale from firm A to firm B for £100 ^a	£20	£20	£0
Sale from firm B to firm C for £300 ^a	£60	£60	£0
Sale from firm C to consumer for £500 ^a	£100	£0	£100
<i>Analysis of firms</i>			
Firm A	£20	£0	£20
Firm B	£60	£20	£40
Firm C	£100	£60	£40

^a Price excluding VAT, which is shown separately in the next column.

The bottom panel of Table 7.1 shows the VAT remitted by each firm. Firm A makes sales of £100 plus VAT with no purchased inputs in our example, so it simply remits the £20 VAT on the sale. Firms B and C each add £200 to the value of the goods, and so each has a net liability of 20% of that (£40). B transforms £100-worth of inputs into £300-worth of outputs, so deducts £20 input VAT (20% of £100) from its £60 output VAT (20% of £300), remitting a total of £40. C is liable for £100 output VAT on its £500 sales, less £60 input VAT on its £300 purchases, also remitting £40 in total. Each firm pays 20% tax on the value it adds; in other words, responsibility for remitting the £100 tax on the total value of the final product is divided across the supply chain in proportion to the value added at each stage.

An alternative, which achieves the objective of taxing only final consumption but without dividing liability across the supply chain in this way, is to make a legal distinction between wholesale and retail sales and tax only the latter. This is the approach of the retail sales tax (RST) which currently operates in most states of the US. This would appear more straightforward, and means that only firms selling to retail customers (firm C in our example) need face the cost of complying with the tax. But the RST approach suffers from significant disadvantages.

First, drawing the distinction between wholesale and retail sales is difficult in practice. An RST requires sellers to establish whether their customers will

use their products for business or consumption. But there is little incentive for sellers to draw the distinction correctly, making misclassification and significant loss of revenue possible. In contrast, VAT requires buyers to establish whether they have used their *purchases* for business rather than consumption. Since only registered traders deduct VAT on their purchases, misclassification of purchases as inputs rather than consumption would normally require people to register for VAT and commit outright fraud. Despite taking a less direct approach than the RST to taxing only final consumption, VAT is more likely to be successful in achieving this goal.²

More importantly, dividing VAT liability across all links of the supply chain means that any one trader evading VAT escapes with only the tax due on the value added in that part of the supply chain, not the VAT due on the whole value of the product. This lessens the incentive for traders to attempt evasion. Traders' claims for deduction of input VAT also require an output VAT invoice from their supplier, so traders buying inputs have an incentive to ensure that their supplier invoices the VAT in full (if not necessarily remitting it to the authorities). If the supplier does not do so, the input buyer ends up paying both parties' VAT liabilities—which is undesirable for the input buyer, but at least means that the government gets the revenue it is due. The symmetric invoices—each claim for input VAT can be checked against the supplier's recorded output VAT—also provide a useful audit trail for the government.

For these reasons, a VAT is a very attractive way for governments to raise revenue. Sadly, it does not always operate smoothly in practice. In the rest of this chapter, we look at some of the more important design problems with the UK VAT. We focus on two issues in particular: first, the widespread use of zero-rating and exemptions, the latter of which in particular causes significant deviations from the 'ideal' described above; and second, the scope

² Distinguishing between business expenditure and consumption expenditure is not always straightforward under a VAT, as e.g. when a firm buys gym memberships for its employees or when a self-employed person buys a computer for personal as well as business use. These blurred borderlines—unlike the more mundane monitoring problem discussed in the text—create difficulties for VAT and RST equally, and indeed for other taxes too: there are close parallels between these boundary issues and the difficulties in identifying work-related expenses for income tax purposes, as discussed in Section 3.2.3.

for non-compliance and the closely related question of how to implement VAT in an international context.

7.2. ZERO RATES, REDUCED RATES, AND EXEMPTIONS

In practice, many goods and services are not subject to VAT at the standard rate (20% from January 2011) in the UK. Some are zero rated, some are subject to a reduced rate of 5%, and some are exempt. The distinction between zero-rating and exemption is that zero-rating allows registered traders to reclaim the VAT on any inputs used in the production process. As a consequence, there is no component of taxation in the final price of a product that is zero rated. Goods and services that are exempt are not subject to VAT when sold, but the producer of an exempt product cannot reclaim the VAT paid on purchases of inputs. The VAT on inputs means that the sale price does include a component of taxation, so is higher than it would be with zero-rating. If a good or service is zero rated, then it is subject to VAT, but the VAT rate is 0%. If it is exempt, then its production is in effect ignored completely for VAT, with no VAT charged on sales or credited on inputs.

Table 7.2 provides a brief summary of the main goods and services that are zero rated for VAT, those facing a reduced (5%) rate, and those that are exempt, alongside government estimates of the revenue forgone by not charging VAT at the full rate on these goods and services (though note that the full rate was 17.5% and 20% during different parts of the year in question). While this table may look relatively straightforward, the reality is that there is a huge amount of detail and complexity in the rules determining exactly what qualifies as, for example, an exempt financial service, zero-rated food, and so on.³

³ Note that Table 7.2 excludes public sector bodies (and others) that are outside the scope of VAT—a status equivalent in effect to exemption. We discuss the treatment of public sector bodies below, but the net revenue implications are small since it mainly affects how much VAT is paid by one part of government to another.

Table 7.2. Estimated revenue cost of zero-rating, reduced-rating, and exempting goods and services, 2010–11

	Estimated cost (£m)
Zero rated:	
Food	14,250
Construction of new dwellings ^a	5,400
Domestic passenger transport	3,250
International passenger transport ^a	150
Books, newspapers, and magazines	1,600
Children's clothing	1,300
Water and sewerage services	1,700
Drugs and supplies on prescription	1,850
Supplies to charities ^a	200
Certain ships and aircraft	550
Vehicles and other supplies to people with disabilities	450
Reduced rated:	
Domestic fuel and power	4,250
Women's sanitary products	50
Energy-saving materials	50
Residential conversions and renovations	200
Exempt:	
Rent on domestic dwellings ^a	4,850
Supplies of commercial property ^a	350
Education ^a	1,150
Health services ^a	1,500
Postal services	200
Burial and cremation	150
Finance and insurance ^a	9,050
Betting, gaming, and lottery duties ^a	1,150
Businesses below registration threshold ^a	1,650

^a Figures for these categories are subject to a wide margin of error.

Notes: These figures refer to 2010–11. VAT was levied at 17.5% during most of that year, rising to 20% for the last three months. With a standard rate of 20% throughout the year, the costs are likely to be around 10–15% higher. Some zero-rated goods (e.g. cycle helmets), reduced-rated goods (e.g. contraceptives, smoking cessation products, and children's car seats), and exempt goods (e.g. cultural admissions charges), where costs are very low or unknown, are not included here.

Source: HMRC statistics, table 1.5 (http://www.hmrc.gov.uk/stats/tax_expenditures/table1-5.pdf).

The overall extent of deviations from uniformity, along with the degree of non-compliance (discussed in Section 7.3 below), can be measured by the ratio of actual VAT revenues to the VAT that would have been levied if all private consumption were successfully taxed at the standard VAT rate. The OECD estimates this ratio at 46% for the UK in 2008, well below the (unweighted) OECD average of 58%.⁴

The list of zero- and reduced-rated goods is extensive. It exists as a result of a combination of distributional concerns and a desire to encourage consumption of particular goods.⁵ The reduced rate on domestic fuel illustrates where these two goals may conflict. It is there for distributional reasons despite the fact that for environmental reasons we might want to discourage, rather than encourage, the consumption of domestic fuel.⁶ In the last chapter, we noted the practical disadvantages of departing extensively from a uniform VAT rate, and the weakness of distributional arguments for zero- and reduced-rating when there are flexible income-related taxes and benefits which can achieve redistribution more precisely and more efficiently. A detailed practical illustration of that is the subject of Chapter 9. The fact that children's clothes are taxed at 0%, children's car seats at 5%, and educational toys (and clothes for children who happen to be big enough to need adult sizes) at 20% is just one of many possible examples of the difficulty in rationalizing the list as it currently appears.

To illustrate specifically how zero-rating works, let us return to our previous example. If firm C's output is zero rated, it need no longer add £100 VAT to the £500 price of its sales, but it can still reclaim the £60 VAT charged on its purchase from firm B. Thus C in effect reclaims the VAT remitted further up the supply chain (£20 by A and £40 by B) and charges no VAT on its own sales. Production in this case is entirely VAT free and the consumer pays a price unaffected by VAT. However, as we shall see in the international context in Section 7.4, the fact that there are firms such as C, which can claim significant net refunds from HMRC as a result of zero-rating, can create opportunities for fraud.

⁴ OECD, 2011.

⁵ The large category of construction of new dwellings does not readily fit either of these explanations. We consider it in more detail in Chapter 16.

⁶ We explore this in more detail in Chapter 11.

Exemption is different. It means that sales are not subject to VAT but, in contrast to zero-rating, the firm cannot reclaim the VAT paid on its inputs. If firm C is selling VAT-exempt goods, it would charge no VAT on its sales but would not be able to reclaim the £60 VAT paid on the inputs it purchased from firm B. Although C's outputs no longer bear VAT, its production costs are now £60 higher, perhaps passed on in a higher price to the consumer. With a final product worth £500, this £60 irrecoverable input VAT represents an effective tax rate of 12%. It should be clear, then, that the effective rate of VAT on the final product depends on the proportion of total value that is added before the exempt link of the supply chain.⁷ The effective VAT rate will always be below the standard rate, but by differing amounts depending on the structure of the supply chain.

Although exempt goods and services bear less than the full rate of VAT, exemption is very different from a reduced rate of VAT. For one thing, exemption is not always more generous than taxation. Where exempt goods and services are sold directly to final consumers, this lower effective rate of VAT is payable *instead of* the standard VAT rate on those sales. But where exempt products are sold to other VAT-registered businesses, the irrecoverable input VAT comes *on top of* the VAT that will be charged on sales to final consumers by businesses further down the supply chain.

If, in our example, it were firm B's output that was exempt from VAT, the £20 VAT on A's sale to B would now be irrecoverable. Crucially, the fact that £60 VAT would no longer be levied on B's sale to C is irrelevant since C could have recovered it anyway; and C's sale to final consumers would be subject to VAT. So the overall VAT payable on this chain of production would be increased by the £20 irrecoverable input VAT; coming on top of the £100 already due on C's sale to final consumers, this means that the final product bears *more than* the full VAT rate.⁸

⁷ Specifically, the effective rate of VAT as a fraction of the full rate is equal to the share of value added before the exempt link. In our example, the goods are worth £300 by the time C acquires them, and the final product is worth £500, so the effective VAT rate is 60% ($300 \div 500$) of the full 20% rate, or 12%.

⁸ In this case, the share of value that is added before the exempt link in the supply chain is 20% (the £100 value of B's inputs is 20% of the £500 value of the final product), so the effective tax rate generated by the irrecoverable input VAT is 20% of the standard 20% rate, 4%. Coming on top of the standard 20% VAT charged on the sale to final consumers, this makes a total

Whether exemption is more or less generous than applying the standard rate thus depends on whether the exempt products are sold to final consumers—in which case the lack of output VAT outweighs the irrecoverable input VAT—or to other businesses—in which case any output VAT would have been recoverable anyway, so the irrecoverable input VAT is a pure extra cost.

Exemption is anathema to the logic of the VAT. It breaks the chain of tax and offsetting credit, leading to distortions of production patterns since taxes on produced inputs cannot be reclaimed. In Section 6.1.1, we stressed the importance of production efficiency and that intermediate inputs to production should not be taxed. The Australian description of exempt activities as ‘input-taxed’ is a good one and immediately draws attention to the inefficiencies that can be created.

We noted above that the effective tax rate entailed by exemption is related to the share of total value that is added before the exempt link in the supply chain. But this share is not fixed, so there is an incentive to minimize it. Exemption creates an incentive to ‘self-supply’—that is, it encourages firms producing VAT-exempt outputs to undertake as many links of the supply chain as they can themselves to ensure that value added at intermediate stages is not taxed. So, for example, firms whose outputs are VAT exempt have a strong incentive to supply their own security services, technical support, cleaning services, and so on, rather than contract them out and face irrecoverable VAT bills. Exemption can create distortions in competition when exempt firms compete with non-exempt firms—favouring exempt over non-exempt firms when selling to final consumers, and favouring non-exempt over exempt firms when selling to other traders—or when competing exempt firms in different EU countries face different costs as a consequence of being charged different VAT rates on their inputs.

Finally, exemption can create additional administration and compliance burdens (and opportunities for tax avoidance) through the need to allocate

effective VAT rate of 24% on the £500 pre-tax price of the final product: £100 output VAT and £20 irrecoverable input VAT making £120 in total.

input VAT between taxable and exempt outputs (credit being available for the former but not the latter) for producers selling both.

While the total cost of these complexities and distortions is hard to ascertain, it is likely to be substantial. Indeed, Maurice Lauré, nicknamed ‘father of the VAT’ for developing the first fully fledged VAT system (introduced in France in 1954), went so far as to describe exemption as ‘the cancer of the VAT system’.⁹ Given this, the natural question to ask is: why is it used? One simple, if unhelpful, answer is that most of the UK’s exemptions, including those for financial services and for health and education services, are mandated by EU rules. The UK government says that ‘A number of goods and services are exempt from VAT because it is considered inappropriate to tax them (including public services such as health, education and welfare) or they are too technically difficult to tax (including financial services)’.¹⁰

We come back to the issue of financial services—important because of its sheer size as well as the unusual problems it poses for VAT—in the next chapter. For other exemptions, the ‘inappropriateness’ referred to in this quote appears to imply some combination of concern about distribution and a view that it is somehow obviously wrong to charge VAT on public services. If the exemptions reflect distributional concerns, then, given the additional distortions created by exemption, our arguments for using other parts of the tax and benefit system to address these apply with even more force than they do to zero-rating. And even if there is some other compelling reason why public services should be treated preferentially, it is far from clear why this preferential treatment should take the form of exemption, which, as we have argued, is far more damaging than, say, applying a zero or reduced rate.

The exemption for services in the public interest, such as health, education, postal, and cultural services, is closely related to the effective exemption applied to many public sector bodies. But the two are not the same, and the relationship between them is becoming more important and more complicated as various forms of privatization, liberalization, outsourcing, and public–private partnerships increase private sector involvement in the provision of public services and blur the boundaries between the two sectors. To some extent, what we have is a VAT system which has just not adapted

⁹ Cited in European Commission (2010c, 28).

¹⁰ HM Treasury and HM Revenue and Customs, 2010, para. 4.2.

with the economy. Blurred boundaries between public and private sectors lead to arbitrary differences in the tax treatment of similar organizations doing similar things. If public and private sector bodies are competing, they might not do so on a level playing field: public sector bodies may have an advantage in providing services to final consumers or to other exempt bodies because of the lack of output tax, whereas private firms' ability to recover input VAT may give them an advantage in providing services to taxable firms.

When public sector bodies are selling something—whether or not in competition with private firms—it might be thought that charging VAT is equivalent to simply adjusting the price: after all, passing the VAT on to HMRC is merely a transfer from one government agency to another, which could (at least in principle) be offset by adjusting the funding of the agency concerned. However, this is not always true: if a public sector body is selling something (for example, parking spaces) that is used both for consumption and as a business input, then charging VAT is not equivalent to a price adjustment as businesses could reclaim the VAT whereas households (and, indeed, exempt bodies paying for the parking space) could not. Since only final consumption should be taxed, prices ought to be higher for households than for businesses. The VAT mechanism achieves this, whereas a simple price adjustment cannot.

Finally, regardless of whether their outputs are sold or provided free to users, exempt public sector bodies have an incentive to self-supply rather than purchase taxed goods and services from private sector suppliers.

In all these cases, distortions could be avoided if VAT were applied to the public sector.¹¹ Studies have shown how this could be done and the benefits it would have,¹² and Australia and New Zealand provide practical examples.

There is one important exemption, though, that is more readily defensible—that for firms (including self-employed individuals) with sales

¹¹ Rules are in place at both EU and UK levels to mitigate some of the distortions, in effect by moving away from exemption. For example, some activities of public sector bodies are classed as 'business activities' and taxed; exemption does not apply where it is deemed that it would lead to a significant distortion of competition; and a special scheme refunds input VAT to some public sector bodies (notably local authorities). However, the scope of these provisions is far from complete and, in practice, many distortions remain.

¹² Aujean, Jenkins, and Poddar, 1999; Gendron, 2005; Copenhagen Economics and KPMG, 2011.

below the VAT registration threshold (£70,000 per year in 2010–11). Registration for VAT is optional for traders with sales below this level. Firms choosing not to register do not remit VAT on their sales, nor can they reclaim VAT on their input purchases, so they are in effect VAT exempt. This is in fact the position of most UK businesses—the government estimates that 2.9 million small businesses are not registered, compared with a total of 1.95 million businesses registered for VAT¹³—although since they are, by definition, small, these unregistered businesses account for only a small minority of sales and revenue. However, many firms with turnover below the threshold choose to register, because if they don't they cannot reclaim VAT paid on inputs. For firms selling mostly to registered traders, any output VAT charged is unimportant because their customers can reclaim it anyway, whereas irrecoverable input VAT could be a significant extra cost. So voluntary registration can often make sense for such firms—although they must also take account of the compliance costs entailed by being registered for VAT.

It is these compliance costs, and the corresponding administrative costs to government, that provide the rationale for a threshold of this kind. The costs of ascertaining VAT liabilities, record-keeping, and so on are substantial, and particularly important for small businesses since many of these costs are fixed rather than proportional to turnover, while the revenue at stake is small. A trade-off needs to be made between the administration and compliance costs of imposing VAT on small businesses, on the one hand, and the loss of VAT revenue and distortion of production activities created by exempting firms below a threshold, on the other. Exempting small firms has all the downsides of exemption generally, discussed above, and the threshold itself brings additional distortions, including creating an incentive for traders to remain below the threshold and giving retailers below the threshold an unfair competitive advantage over taxed retailers. These costs and benefits are difficult to quantify. Crawford, Keen, and Smith (2010) provide some indicative calculations that weigh administrative and compliance costs against lost revenue (but ignore distortions). The results prove highly sensitive to difficult-to-measure parameters, but on balance we

¹³ HM Treasury and HM Revenue and Customs, 2010, para. 4.9.

agree with them that ‘there is good reason to suppose that the relatively high threshold should be counted as a strength of the UK VAT’.

Another route, used in many countries, to minimizing administration and compliance costs is to apply a simplified scheme to small businesses. Such a scheme has existed in the UK since 2002 for small firms (those with non-exempt sales below £150,000, excluding VAT, in 2010–11), which have had the option of using a simplified flat-rate VAT scheme. Under the flat-rate scheme, firms pay VAT at a single rate on their total sales and give up the right to reclaim VAT on inputs. The flat rate, which varies between 4% and 14.5% depending on the industry,¹⁴ is intended to reflect the average VAT rate in that industry after taking into account recovery of VAT on inputs, zero- or reduced-rating of some outputs, and so on. This scheme has problems. By disallowing the recovery of VAT on inputs, it distorts production decisions in the same way as exemption. The differentiation of rates between 55 categories of industry creates distortions and policing problems at the boundaries between them. And the scheme may not even succeed on its own terms, in reducing compliance costs for relatively small businesses. It is not clear how much easier it is to comply with the flat-rate scheme than with the standard scheme; more importantly, optional schemes of this kind naturally encourage firms to estimate (at least roughly) their liability under both regimes to see which is lower—indeed, tax advisers often insist on doing both sets of calculations for fear of being found negligent if clients choose the wrong option. This increases compliance costs—ironic when the scheme exists precisely to reduce them—as well as ensuring the maximum revenue loss for the government.

While the merits of an optional flat-rate scheme are questionable, the case for a substantial registration threshold is strong. In other areas, though, there is a powerful case for reform. In general, we agree with the view expressed by Crawford, Keen, and Smith (2010, 301) that ‘the extensive rate differentiation still found in the VATs of EU members is coming to look increasingly quaint’. It reflects an outdated view of what it is possible to do within a VAT system and does not reflect lessons learned by those countries, including Australia and New Zealand, that introduced VATs more recently and with many fewer exemptions and less rate differentiation than are seen

¹⁴ This is the range of rates from January 2011, when the main VAT rate rises to 20%.

either in the UK or in many other European countries. The UK's long list of exemptions and extensive zero-rating are increasingly hard to justify. Some of the changes needed are substantial and complex and we don't deal with them all in this chapter. Rather, we devote Chapter 9 to illustrating how the use of zero and reduced rates of VAT can be drastically reduced without adverse consequences for either redistribution or work incentives. We discuss in detail in Chapter 8 how the major issue of the exemption of financial services might be dealt with. And the treatment of housing is dealt with in Chapter 16. For the rest of this chapter, we focus first on some of the general compliance issues that are thrown up by the design of VAT and then on the international context and particular compliance issues surrounding the treatment of exports.

7.3. NON-COMPLIANCE

Evasion and fraud are important issues in the administration of VAT. Keeping up with what remains a vast paper trail of invoices is a formidable task. HMRC estimates that the 'VAT gap' was £11.5 billion in 2009–10.¹⁵ That is the difference between tax actually collected and the tax that would have been paid if all individuals and companies complied with both the letter of the law and HMRC's interpretation of the intention of Parliament in setting law. At 14% of the potential revenue yield—higher than for most other taxes—it is clearly a cause for concern.

Not all of the VAT gap represents outright fraud: a significant part of it reflects innocent error or legal tax avoidance, for example. But illegal evasion is significant. In general terms, evasion falls into two main categories:

- traders understating taxable sales and/or overstating creditable inputs;
- traders disappearing without paying a VAT bill they owe.

The first category involves a range of different practices. These include working cash-in-hand and not recording sales that ought to be taxable, or failing to register for VAT despite being liable. Invoices for input purchases

¹⁵ HM Revenue and Customs, 2010d, para. 1.19.

can be faked, or it is possible to claim that sales are zero rated (for example, by faking export invoices) when they should not be. Evaders can also exploit the different rates of VAT on different forms of transaction, taking advantage of the difficulty in policing borderlines between different activities (for example, consumption versus business expenditure; inputs to exempt versus non-exempt activities; inputs from registered versus unregistered suppliers; taxable versus zero-rated inputs). Some of these problems are inherent to a VAT system, though many are concrete and expensive examples of the consequences of the complexity created by deviations from uniformity. The way VAT works does limit the scope for evasion because it is harder to understate sales when the buyer wants an invoice with which to reclaim input VAT and, correspondingly, it is harder to overstate inputs when one needs an invoice from the seller. Broadening the VAT base would further help, since reducing the number of boundaries would leave less scope for misclassification—reducing error and avoidance as well as evasion. It would also be harder to claim zero-rated sales if fewer products were zero rated. Other aspects of VAT policy—such as the choice of registration threshold, the speed with which payment is demanded and refunds are given, and the sheer level of resources devoted to HMRC’s enforcement activities—could also have an impact on evasion, though of course there are also other considerations involved in each of these choices.

The second form of evasion mainly arises when individual traders have large net VAT liabilities. The fractional nature of VAT is designed precisely to deal with this problem: the VAT liability on a final consumption sale is divided across the supply chain so that no individual trader gains that much by disappearing. Of course, where a single trader genuinely creates significant value added, there is still a substantial incentive to disappear, but much less so than under a retail sales tax. And the very fact that the value added is genuine must reduce the incentive to sacrifice the long-term benefits of remaining active for short-term fraudulent gains.

Those traders with the biggest incentive to evade VAT in this way are those with large liabilities relative to their turnover. These will generally be firms that produce taxed outputs using untaxed inputs. There is not much scope for this in a purely domestic context because most zero-rated items are final consumption goods. But, as we shall see, the zero-rating of exports does create significant possibilities for fraud in an international context.

7.4. THE INTERNATIONAL CONTEXT

So far, we have considered the implementation of VAT in a purely national context. But the international dimension, and particularly the EU dimension, is central to the operation of VAT. For one thing, the EU is a major player in VAT policy—indeed, the UK adopted a VAT in 1973 largely because it was a precondition for entry to what was then the European Economic Community. As well as setting out standardized definitions and rules, the EU mandates a minimum standard rate of 15%, restricts the use of reduced rates, forbids the extension of zero-rating to new items, and insists on various exemptions. But significant changes to the economic environment in recent years have brought international tax issues ever more to the fore. For EU member states, the most important of these was the completion of the single European market in January 1993, which directly affected the operation of the VAT system. The completion of the single market also contributed to the wider impact of globalization in increasing trade. Particularly marked have been the very large increase in cross-border trade in services and the birth and rapid expansion of e-commerce, both more difficult to tax than traditional trade in physical goods. These changes have highlighted weaknesses in the current tax system and constrain the design of an alternative system.

In an international context, there is a fundamental question over *where* taxation should take place—in the country in which consumption takes place (the *destination principle*) or in the country of production (the *origin principle*).

It is worth clarifying what taxation ‘in the country of consumption or production means. First of all, while consumption generally takes place at a single location, production of a good or service is often split between several countries. Under origin taxation, the final price paid by the consumer would include some VAT at the rates charged in each country, in proportion to where the value was added. But, as we shall see, difficulties in identifying where the value was really added can cause problems.

Second, the country whose tax rate determines the final burden of tax on a product may not be the country that initially collects the revenue or the country that ultimately receives the revenue. In line with the economic literature (but unlike much EU policy discussion in recent years), we use

origin and destination taxation to refer to the first of these three. But while current international practice is for VAT to be destination-based in all these senses, some reform proposals involve separating these different aspects, and terminology often becomes confused.

A uniform tax levied on a destination basis is a tax on the value of aggregate consumption, while a uniform tax levied on an origin basis is a tax on the value of aggregate production. In principle, there need not be any stark contrast between the two. When trade is balanced—as ultimately it must be—these are the same thing. Moving from a uniform destination-based VAT to a uniform origin-based VAT would leave trade patterns and economic welfare unchanged as exchange rates and/or prices in the different countries would adjust to offset any impact on the price of each country's products to purchasers in other countries.

Reality is not so neat.¹⁶ In practice, VAT systems are far from uniform, so an across-the-board adjustment to a country's exchange rate or aggregate price level could not be a perfect offset for all goods.¹⁷ And, to be neutral, a shift to the origin basis would have to be applied—unilaterally if necessary—to trade with all countries, which is unlikely to be politically popular because it would give the appearance (in fact illusory) of making exports uncompetitive on international markets, and would probably be inconsistent with World Trade Organization agreements. Thus origin and destination bases for taxation are not fully equivalent in practice; but what these arguments bring home is that the considerations involved in choosing between them are subtler than it might first appear.

The standard international practice of using a destination basis has the advantage that businesses and consumers are indifferent between imports and domestically produced goods, and between imports from different countries that levy different tax rates. In both cases, the only tax payable is that imposed in the country of consumption. This means that it should not distort decisions over where to produce.

The most problematic aspect of the destination basis is how it can be implemented in the context of an EU without internal border controls. The

¹⁶ The conditions for equivalence are set out in Lockwood, de Meza, and Myles (1994).

¹⁷ The precise condition for equivalence is in fact slightly weaker than full uniformity, but it is still far from being met.

problem arises for both sales to other businesses and cross-border shopping by individuals.

At present, the VAT system ensures that goods are taxed only in the country of consumption by zero-rating exports (thus freeing them of all VAT levied on the supply chain up to that point) and subjecting all imports to tax. In the EU prior to 1993, this involved the use of border controls to monitor exports and imports. Since then, border controls within the EU have been abolished to facilitate free trade and create a level playing field between firms operating across member states. But this means that it is no longer possible to apply at borders the tax adjustments that are fundamental to the operation of the destination principle. We discuss the problems this creates, and possible ways to deal with them, below.

As far as individual cross-border shopping is concerned, if two countries set different tax rates on a product, then, when there are no borders (and so no way to enforce limits on what individuals can bring home from abroad having paid tax only at the foreign-country rate), consumers can purchase the product in the country with the lower rate of tax and ship it home. This form of cross-border shopping is clearly an inefficient outcome. The origin principle does not suffer from this disadvantage: since tax liability depends on where the goods are produced rather than where they are consumed, consumer prices will tend to be equalized across countries, which would not distort where goods are consumed. There is simply no need to go abroad to buy goods more cheaply if the tax rate is the same wherever they are purchased. However, an origin-based tax would affect the pattern of production. Whether it is more important to avoid distortions to production patterns or to consumption patterns is still an issue for debate.¹⁸ But an important downside of the origin approach is that it encourages producers to use mechanisms such as transfer prices (the prices at which part of a firm in one country 'sells' its output to another part of the firm in another country) to lower their tax bills artificially. Since, as mentioned above, pure origin taxation involves taxing the value added in each country at that country's tax rate, firms can use transfer prices to shift measured value added into

¹⁸ Key contributions to the debate include Lockwood (1993), Keen and Lahiri (1998), Haufler and Pflüger (2004), Keen and Wildasin (2004), Haufler, Schjelderup, and Stahler (2005), and Hashimzade, Khodavaisi, and Myles (2005). Crawford, Keen, and Smith (2010) provide a summary of the arguments.

jurisdictions with low tax rates, just as they have an incentive to use transfer pricing so that profits appear to be earned in jurisdictions with low corporation tax rates. The problems associated with transfer pricing in corporation tax suggest that this is a potentially major drawback.

Since corporation tax is levied essentially on an origin basis (the usual terminology is 'source basis' in that context, but there is little difference), the pros and cons of origin and destination principles are relevant not just for the design of the VAT system itself, but also for governments deciding how much to rely on direct (corporate and personal) versus indirect taxes. At present, there is little prospect of the EU (let alone the UK unilaterally) moving to a pure origin basis for VAT, or for that matter a destination basis for corporation tax.¹⁹ From the point of view of the UK (and the rest of Europe) in 2011, the most important international issue in VAT implementation is that of how best to manage a destination-based system in the absence of border controls. This has confronted the operation of VAT in the EU with some significant challenges, including frauds that have exploited the zero-rating of exports. We look now at this issue in more detail.

7.4.1. Export Zero-Rating and Compliance

We have already seen that zero-rating leads to opportunities for evasion. Zero-rating of goods at export involves tax authorities paying out large-scale refunds to exporting companies, since there is no tax on their sales and they can reclaim input VAT on their purchases. If this reclaim is accompanied by fraud or failure to pay VAT further down the chain, then not only can the revenue authorities collect less tax than intended, they can even end up paying more in refunds than they collect in tax.

In the early and middle years of the 2000s, missing trader intra-community (MTIC) fraud, including so-called carousel fraud, which exploited this possibility, became a widespread problem across the EU. In 2004, the European Commission reported that losses from fraud, of which carousel fraud is the best-known (but not the only) manifestation, amounted

¹⁹ Auerbach, Devereux, and Simpson (2010), however, argue that shifting corporation tax to a destination basis deserves serious consideration. The international dimension of corporation tax is the subject of Chapter 18.

to 10% of net VAT receipts in some member states. One result for the UK was that, in the middle years of the 2000s, the Treasury was consistently forecasting VAT receipts several billions of pounds greater than actual out-turns because of losses to fraud. HMRC estimates that MTIC fraud in 2005–06 is likely to cost the Exchequer between £2.5 and £3.5 billion.²⁰ This was enough to distort trade statistics. The Office for National Statistics estimates that there were a staggering £20.7 billion of trade flows associated with MTIC fraud in the first half of 2006 alone²¹—though that appears to have been the peak. For an otherwise rather abstruse part of the tax system, these issues have certainly received a remarkable amount of press coverage,²² in part because of the large totals involved and in part because some spectacular cases appear to have made some individuals very rich very quickly.

Whilst the scale of this fraud appears to have dropped dramatically in recent years, to between £0.5 and £1.5 billion in 2009–10,²³ it is worth explaining how it works since it provides a very clear illustration of some of the weaknesses in the VAT system, and in particular why the issue of the appropriate treatment of exports for VAT purposes has been the subject of considerable attention.

A carousel fraud is operated by importers purchasing products that are zero rated and selling them on with VAT added to another trader. The purchasing trader reclaims the input VAT, but the seller does not pay the VAT due and disappears. The way this works is illustrated in Figure 7.1. The importing company (company B) pays no VAT on its purchase because the import is zero rated. It then sells the goods to another company (company C), legitimately charging VAT on the sale. It should then remit this VAT to the tax authorities. Company C may well be an entirely innocent ‘buffer’ in the fraud. In the simplest version, it then sells goods on to company D, again charging VAT. Company D then exports the goods back to the original company, A, charging no VAT on the sale since exports are zero rated, but claiming a refund of input VAT on its purchase of the exported goods.

²⁰ HM Revenue and Customs, 2010d, table 2.4.

²¹ Page 8 of <http://www.statistics.gov.uk/pdfdir/trd1007.pdf>.

²² For example, ‘Revealed: the £5bn-a-year tax fraud’, *The Guardian*, 9 May 2006. The BBC *Panorama* programme on 16 July 2006, ‘Do you want to be a millionaire?’, also investigated carousel fraud; details at <http://news.bbc.co.uk/1/hi/programmes/panorama/5366914.stm>.

²³ HM Revenue and Customs, 2010d, table 2.4.

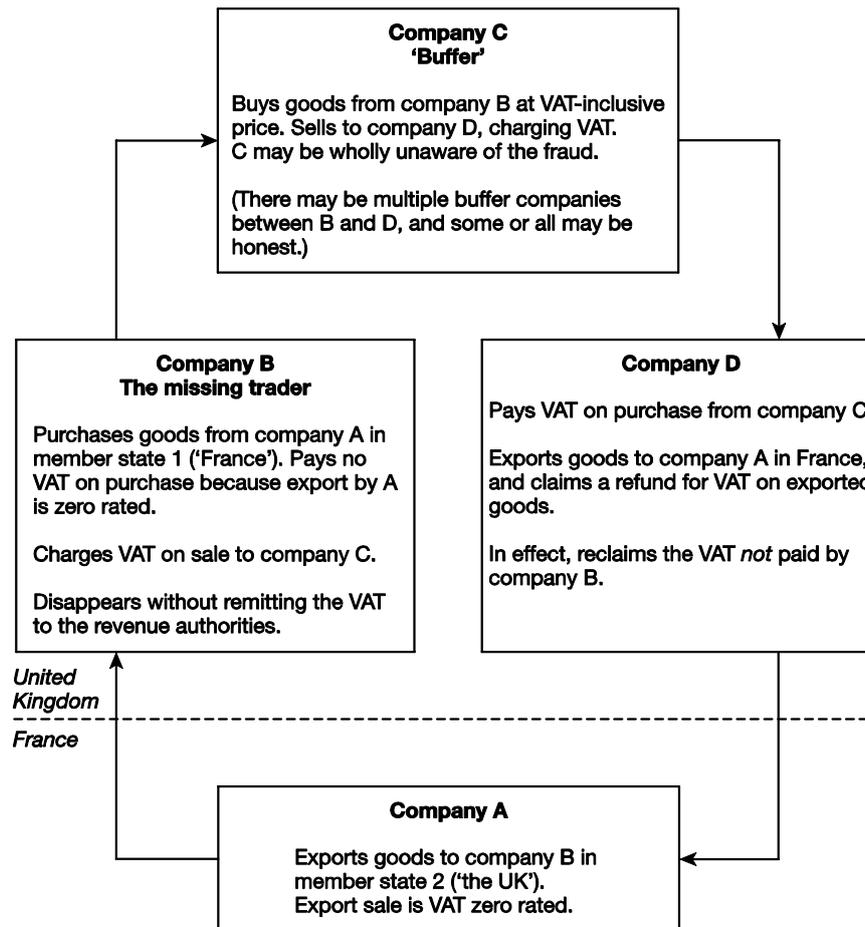


Figure 7.1. A simple illustration of carousel fraud

Source: Crawford, Keen, and Smith, 2010.

This is all fine unless the original importing company, B, 'disappears' before remitting VAT to the authorities. If this happens, then, in effect, company D is claiming a refund for VAT which was never actually paid. If there were no zero-rating of exports, then the importing company would have been charged VAT by the original exporter, which it could not then reclaim as input VAT if it disappeared without remitting its output VAT. And the final exporter would not be entitled to any refund of VAT if its sales were not zero rated. The opportunity for this type of fraud would not exist.

7.4.2. Responses to the Break in the VAT Chain at Export

Fundamentally, it is the break in the VAT chain created by the zero-rating of exports which allows this type of fraud to occur. It is particularly problematic for tax authorities because it involves not just reduced revenue but also an actual payment out from the system—revenue is negative. It has been a particular issue within the EU since completion of the single European market in January 1993. Before 1993, the border controls between member states permitted a destination system to operate, with commodities taken out of tax in the exporting country and brought into tax in the importing country. This was achieved by the zero-rating of exports, combined with subjecting all imports to VAT at the border. Since 1993, border controls within the EU have been abolished. So the ability to raise tax at the border has been lost, and the system depends upon a paper trail of invoices—account auditing plays the role previously performed by frontier controls. This has undermined the operation of the destination system and opened up the potential for abuse. Imports from outside the EU are, of course, still captured at borders.

The government's preferred solution to MTIC fraud has been *reverse charging*. In business-to-business transactions, this places the VAT liability on the buyer rather than the seller. This would deal effectively with the carousel fraud illustrated, because the VAT due on the sale by company B (the missing trader) would become the responsibility of the buyer, C. In turn, the tax due on the sale from C to D would be the responsibility of D. The zero-rating of the subsequent export sale would then offset D's tax liability on its purchases from C, reducing the tax payment by D but not requiring outright refunds. The opportunity to make fraudulent gains by claiming refunds of tax that has not in fact been paid would thereby be eliminated.

So reverse charging reduces the possibility of fraud by preventing the purchaser claiming back input VAT which has not actually been paid. But it also fundamentally undermines the fractional nature of the VAT, which we have argued is a crucial part of its attraction. If reverse charging applies all the way down the line, then nothing is collected until the final transaction and we effectively have a retail sales tax. Nevertheless, the UK has implemented reverse charging for mobile phones, computer chips, and some

other easily transported and high-value goods which have been particularly prone to carousel fraud. Other member states (Austria and Germany, for example) have made applications for more comprehensive systems of reverse charging, but these applications have been rejected by the European Commission.

Reverse charging does appear to have met with some success, at least in the short term, in reducing MTIC fraud. But by undermining the fractional nature of VAT collection and creating further distinctions between those goods and services that are subject to it and those that are not, it potentially opens up the route to more mundane forms of evasion. It also leaves open the possibility that the carousel fraud itself could move to other goods which are not subject to reverse charging. It looks more like an interim solution plugging a hole than a fundamentally stable solution.

To understand what the more fundamental reform options are, recall that zero-rating of exports has the following features. All VAT in the exporting country is 'washed out' of the good at the point of export—the exporter receives a VAT refund. The good is then brought into the VAT system of the importing country. All VAT on the final product is paid at the rate of the country where consumption occurs and all revenues are collected by, and enjoyed by, the country where consumption occurs. This is a pure destination system.

The most radical option would be to move to a pure *origin* system, the advantages and disadvantages of which we discussed earlier. In practical terms, this would involve exporters remitting VAT on their sales at the exporting country's VAT rate (rather than zero), while importers would claim credit from the importing country's tax authorities at the destination country's VAT rate, irrespective of the VAT charged in the country of export. Note that this would mean not only that the burden of VAT on the final product reflected the rate in the country of production, but also that the revenue accrued to that country, so it would involve a redistribution of revenues between countries, as exporting countries would gain revenue and importing countries would lose revenue.

A variant on this is *exporter rating*. Under this system, exports would again carry the VAT of their country of origin, but now it would be this origin-country rate of VAT that could be reclaimed as input VAT in the destination country. Because the VAT could be reclaimed at the original rate charged,

the VAT charged to the final consumer would still be that of the destination country. Exporter rating therefore has the underlying economic characteristics of destination-based taxation—for example, in avoiding the transfer pricing problems that could be introduced by a full origin system.²⁴ And unlike zero-rating of exports, it would avoid breaking the VAT chain at export. But it would create its own problems, most important of which would be the incentives that would be created for countries to impose high VAT rates on goods that are primarily exported. The country of production would collect the VAT. The importing company wouldn't worry about the high VAT rate because it could claim that VAT back in full. But the government in the importing country would clearly lose revenue. This could be avoided by having a 'clearing house' which would reallocate revenues to match revenue collection under the current system. But this would undermine the enforcement incentives for individual countries: tax authorities have little incentive to devote resources to ensuring that exporters remit their full tax liabilities if this revenue is then to be handed over to the importing country's government. Enforcement incentives could be reinstated by fixing flows from the clearing house at some agreed level (perhaps based on statistics for aggregate consumption or trade) rather than transferring the actual VAT collected on exports, but then countries would again be incentivized to increase VAT rates to increase revenues from the taxation of exports. So while a form of exporter rating has long been envisaged as a goal of the European Union,²⁵ no such scheme has been adopted, because of these incentive and administrative difficulties.

An alternative way to avoid the break in the VAT chain is to set a single EU-wide VAT rate (the 'intermediate' rate)—no higher than any current VAT rates—either on all business-to-business transactions or on all cross-border supplies within the EU, with countries free to set their own tax rate

²⁴ Confusingly, however, discussions of exporter rating in the EU context often refer to it as 'taxation at origin', since that is where the revenue would be collected.

²⁵ In 1987, the European Commission proposed that exporter rating be adopted after the abolition of border controls. A subsequent proposal (European Commission, 1996) involved exporter rating based on the place of establishment of the seller rather than the physical movement of goods. Both attempts were unsuccessful, though a formal commitment remains (in the VAT Directive—Council Directive 2006/112/EC of 28 November 2006 on the Common System of Value Added Tax) to replace the current system—still described as 'transitional'—with a 'definitive' regime.

above the EU-wide rate for domestic sales to final consumers if they wish. This system of *uniform rating* has been proposed in a number of variants, each with its own disadvantages.²⁶ It could apply either just to exports (the ‘CVAT’ proposal²⁷), which means sellers would still face different procedures for domestic and cross-border sales, or to all transactions between registered traders (the ‘VIVAT’ proposal²⁸), which would require making a distinction between business-to-business and business-to-consumer sales. It could be operated as a separate EU-level mechanism (as proposed for CVAT), which would require a new EU-level bureaucracy, or be incorporated into the VAT administrations of member states (as proposed for VIVAT and for a variant of CVAT examined by the European Commission in 2008). The latter would require reallocation of revenues to importing countries through a ‘clearing house’ mechanism similar to that envisaged for exporter rating, except that the commonality of the intermediate rate allows formula-based revenue allocations to be used with no scope for countries to respond to tax-setting incentives. But all of these variants would retain the economic properties of destination-based taxation, fix the break in the VAT chain between member states (to an extent that would depend on the level of the EU-wide intermediate rate), and leave incentives for enforcement and rate-setting unaffected, without diminishing the current ability of member states to choose the rate applied to domestic sales to final consumers and hence how much revenue they raise.

7.4.3. Future Directions

Whilst VAT is in many ways a successful tax, the audit trail that is required to collect it accurately and effectively is complex. As we have seen, scope for fraud and evasion is significant, particularly in the context of intra-EU trade, where a combination of zero-rating of exports, the lack of internal frontiers, and a reliance on paper invoices has allowed significant frauds. There are

²⁶ As well as the variants mentioned here, there are others, such as the ‘DVAT’. DVAT is discussed in Crawford, Keen, and Smith (2010), on which this section draws heavily.

²⁷ The ‘compensating VAT’, proposed by Varsano (2000) for Brazil and McLure (1999 and 2000) more generally.

²⁸ The ‘viable integrated VAT’, proposed by Keen and Smith (1996 and 2000).

no easy solutions. Reverse charging has been effective in the short run. Increased auditing has also helped. More fundamental changes have their own problems, creating different sets of problematic incentives and/or requiring additional bureaucracy to keep national VAT revenues in line with current levels. Of the reform options, some form of uniform rating looks the most promising. But it would not be in any sense straightforward to implement.

It is possible that a longer-term solution does exist, though, based on a much more effective enforcement system using new technology. The current system remains heavily reliant on paper invoices. It is very hard to follow the VAT payment trail through the supply chain. There are also significant delays between the point at which firms charge VAT to their customers and the point at which they remit the VAT to the authorities. (Recall that the basic problem in the carousel fraud illustrated in Figure 7.1 was that the importing firm ‘disappeared’ before it had remitted the VAT it owed.)

One increasingly plausible possibility is that developments in electronic payment systems might overcome many of these problems. They could allow—and the authorities could insist upon—direct payments of VAT at the point of any transaction. They could also allow automatic reconciliation between the VAT liabilities of firms at different points in the production chain. The temptation to put one’s faith in new technology often leads to disappointment, but there do seem to be good reasons for thinking that many of the current difficulties in the system, and the bureaucracy around exports, could be overcome in this way. If that were possible, then the balance of advantage between the different possible VAT structures we have been looking at would change. Effectively, the current system of zero-rating of exports could work much more readily—electronic checking and payments would stand in the place of the dismantled border controls.

It may be that the best approach for the next few years will be to use the necessary sticking plasters—such as reverse charging for some goods—to keep the current system operating, while planning a technological overhaul of the administration to put the current system on a stable long-term footing. But we are not best placed to judge the viability of alternative technologies. Unless the view is taken that a technological solution is genuinely feasible and planning begins accordingly, policymakers should

instead look to move away from zero-rating exports, probably towards some form of uniform rating. The status quo cannot hold indefinitely.

7.5. CONCLUSIONS

VAT is an appealing way to raise revenue. In its purest form, it taxes only final consumption. Because it is collected at stages through the supply chain, scope for wholesale evasion, such as can exist with a retail sales tax where all the revenue is collected at the point of sale to the final consumer, is limited. But many difficulties remain which limit its effectiveness in practice. VAT is complex to administer, and depends for its operation on careful auditing and enforcement. Evasion remains a problem. Rate differentiation and the use of exemptions create welfare-reducing distortions as well as adding to complexity.

Some of its difficulties are perhaps largely unavoidable. Trade-offs need to be made over the scale of business that should be brought within the ambit of VAT—we believe a relatively high threshold, such as in the UK, has much to commend it, minimizing as it does compliance costs for small businesses. The appropriate treatment of exports within the EU in the context of no border controls requires one to trade off advantages and disadvantages of different systems—though it is no surprise that the current system of zero-rating of exports has always been considered interim. Interim, however, is looking increasingly permanent.

But there are clear directions for reform that would make the VAT system much better. The first, and most important, is a wholesale removal of most of the exemptions and zero- and reduced-rating of goods and services which add so much complexity and distortion to the current system. The next two chapters look at the specific issue of applying VAT to financial services and the more general issue of how the scope of zero- and reduced-rating can be much reduced whilst maintaining equity and work incentives. It should be possible, and is certainly desirable, to move in this direction even in the case of activities, such as public services, that have traditionally been seen as outside the scope of VAT.

The choice for the treatment of intra-EU trade seems to come down to either continuing the current sticking-plaster solution of selective reverse charging whilst waiting for, or rather actively planning for, a revamp of administration and enforcement through electronic payment systems, or else a fundamental overhaul of the EU system, moving towards a system of uniform rating of the kind we have outlined.