



WP/07/31

IMF Working Paper

VAT Fraud and Evasion: What Do We Know, and What Can be Done?

Michael Keen and Stephen Smith

IMF Working Paper

Fiscal Affairs Department

VAT Fraud and Evasion: What Do We Know, and What Can be Done?

Prepared by Michael Keen and Stephen Smith¹

February 2007

Abstract

This Working Paper should not be reported as representing the views of the IMF.

The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

Like any tax, the VAT is vulnerable to evasion and fraud. But its credit and refund mechanism does offer unique opportunities for abuse, and this has recently become an urgent concern in the European Union (EU). This paper describes the main forms of noncompliance distinctive to a VAT, considers how they can be addressed, and assesses evidence on their extent in high-income countries. While the practical significance of current difficulties in the EU should not be over-stated, administrative measures alone may prove insufficient to deal with them, and a fundamental redesign of the VAT treatment of intra-community trade required. The current difficulties in the EU largely reflect circumstances that would not apply in the United States.

JEL Classification Numbers: H20, H21, H22

Keywords: value-added tax, sales tax, enforcement, compliance, noncompliance, audit

Author's E-Mail Address: mkeen@imf.org; stephen.smith@ucl.ac.uk

¹ Fiscal Affairs Department, International Monetary Fund, Washington, DC 20431, telephone +1 202 623 4442, Email: mkeen@imf.org; Department of Economics, University College London, Gower Street, London WC1E 6BT, United Kingdom, telephone +44 20 7679 5882, Email: stephen.smith@ucl.ac.uk. We are grateful to Bill Fox, Jean-Paul Bodin, Graham Harrison, and Rosanne Altshuler for helpful comments, advice, and suggestions. Views and errors are ours alone, and should not be attributed to the International Monetary Fund, its Executive Board, or its management.

Contents	Page
I. Introduction	3
II. Evasion, Enforcement, and the VAT	5
III. International Aspects of VAT Fraud.....	12
IV. The Extent of VAT Noncompliance and Fraud.....	15
V. Enforcement Strategies	22
VI. Conclusions—with Lessons for the United States.....	28
 Tables	
1. HMRC’s “Top Down” Estimate of the United Kingdom VAT Gap, Tax Years 2000–01 to 2004–05	16
2. Bottom-Up Estimates of VAT Revenue Losses, United Kingdom, Tax Year 2001–02.....	18
3. Estimates of the VAT Gap for Ten EU Countries	19
 Figures	
1. The Basic Carousel Fraud.....	13
2. The VAT Gap in the United Kingdom, 1990–01 to 2004–05	17
 Boxes	
1. The Rudiments of VATs Peak	4
2. Equivalence of VAT and RST—When They Both Function Perfectly	6
References.....	30

I. INTRODUCTION

The key claim made by advocates of the VAT is that it is a particularly effective way of raising tax revenue: Cnossen (1990), for example, argues that “purely from a revenue point of view, VAT is probably the best tax ever invented,”² Advocates have also long recognized, of course, that the VAT, like any other tax, is vulnerable to evasion and fraud, but stress distinctive features of the VAT that may make it less vulnerable than other forms of taxation. Over the last few years, however, there has been a marked increase in concern with losses of VAT revenue through evasion and fraud. This is especially so in the European Union, where the abolition of internal EU frontiers at the end of 1992 opened up new areas of vulnerability. The European Commission (2004), for example, reports that losses from fraud—most famously, “carousel fraud”—have recently amounted to 10 percent of net VAT receipts in some member states. In March 2006 the first ever fall in annual nominal VAT receipts in the United Kingdom was attributed by some to fraud (Ainsworth, 2006), which has even come to seriously distort trade statistics (Ruffles and others, 2003). The spectacular nature of some of these frauds—such as that of the Glasgow businessman who netted over \$8 million in just four months (and wound up with four years in prison)—has also attracted considerable media interest, including lead stories in the national press.³ These difficulties are naturally troubling not only for countries that have a VAT in place but also for those, like the United States, that might consider adopting one.

This paper reviews what is known about the exposure of the VAT to revenue losses through noncompliance, with a particular focus on fraud and evasion,⁴ and considers what can be done in response. The topic is broad, and the coverage here selective. Problems of VAT fraud and evasion are not confined to EU members, or to high income countries more generally. Sri Lanka, for example, has recently lost substantial revenue—reportedly about 10 percent of its net VAT receipts—from a single fraudulent episode.⁵ Indeed, the more pervasive informality in developing countries, and typically higher levels of corruption, suggest that fraud and evasion are likely to be even greater problems there. It is the difficulties faced by higher-income countries, however, that are the focus of attention in this paper, since it is their experience that is likely to be most informative in understanding the risks which any VAT in the United States would face. And among higher-income countries we focus particularly on the EU, where problems of fraud and evasion have risen to the prominence and topicality indicated above. For the obvious question arises: Should the current difficulties in the EU

² Whether or not the VAT has lived up to its reputation as a “money machine” is analyzed by Keen and Lockwood (2006a,b) for, respectively, a wide range of countries and the OECD in particular.

³ “Revealed: the £5bn-a-year tax fraud,” *The Guardian*, May 9, 2006.

⁴ A few words on terms as they are used here. By “noncompliance” we mean any component of the “tax gap” between the amount of tax actually paid and that theoretically due under full compliance. This is somewhat broader than evasion and fraud (the latter differing from the former in involving some positive act of commission), in that payment may be less than full as consequence of genuine error by the taxpayer—or, as will be seen, by the tax authorities.

⁵ *Tax Notes International*, April 17, 2006.

give particular reason for the United States to hesitate when contemplating the introduction of a VAT?

The discussion that follows starts by stressing the importance of enforcement issues to the case for using indirect taxes in general and the VAT in particular. We then describe the main types of fraud and evasion to which the VAT is liable, giving special attention to the risks that arise from the zero-rating of exports—especially the ‘carousel’ frauds that have attracted so much attention, and which epitomize the core problems in this area. Next comes a review of the (quite scant) empirical evidence on the scale of VAT noncompliance (and, briefly, how it compares in this respect with other taxes). The paper then turns to options for tackling the problems of VAT fraud and evasion, again focusing on current concerns in the EU. A final section concludes.

As background for all this, Box 1 reviews some key VAT terms and concepts.

Box 1. The Rudiments of VATs Peak

The defining characteristic of a VAT is that it is, in principle (there are always exceptions), a broad-based tax on all commodity sales, whether to consumers or to other businesses. However, registered businesses are able to credit the VAT charged on their purchases (“*input VAT*”) against the VAT due on their sales (“*output VAT*”)—an example is in Box 2 below. Any excess credits this creates are refunded to the taxpayer.

Zero-rating simply means that VAT is levied at a rate of zero: no tax is due on output, but the credit—which, with no output tax, becomes a refund—is still available for input VAT.

Exemption, in contrast, means that no tax is due on output but nor is a credit available for input VAT. Thus the VAT ‘sticks’ on business purchases; the Australian term “input-taxed” is more evocative.

Under the “*destination principle*”—which is the international norm—commodities are taxed by the jurisdiction in which they are consumed. This is generally implemented under the VAT by zero-rating exports and charging VAT on imports (although, as discussed later, there are other ways of implementing the destination principle). While this implies that exporters will be due refunds, it does not amount to an export subsidy and is fully WTO-consistent: it is simply a device for removing any domestic VAT from exports. Levying VAT on imports simply puts them on the same basis as their domestically produced counterparts.

(The alternative to destination taxation is the *origin principle*, by which commodities are taxed according to where they are produced. Although some countries of the former Soviet Union applied VAT on this basis for some trade flows among themselves in the early years of transition, no country now applies such a system.)

The VAT as defined above can be implemented in a number of ways. The focus throughout this paper is on the “*invoice-credit*” form, under which registered traders charge tax on their sales and issue corresponding invoices to their customers, who, if also registered, can use these invoices to establish a right to credit or refund against their own output VAT liability. All national-level VATs of which we know are of the invoice-credit form, except that Japan applies a *subtraction method* VAT, under which each trader is taxed on the difference between their purchases and sales. Alternative forms of VAT are sometimes applied at lower levels of government: the Italian IRAP is a subtraction method VAT, and Michigan applies—though this is to be removed in 2008—an addition method VAT (levied on the sum of wages and profits).

II. EVASION, ENFORCEMENT, AND THE VAT

This section considers some basic issues of enforcement in relation to commodity taxation in general and the VAT in particular.

Enforcement issues in the design of indirect taxes

The case in pure theory for the use of any kind of general commodity tax is an uneasy one: a uniform tax on consumption is equivalent, in present value, to a uniform tax on wage income (and net bequests), and in that sense adds little to what can be achieved by simply taxing wage income. Atkinson and Stiglitz (1980) show further that if preferences are weakly separable between consumption and leisure, then differential commodity taxation has no distinctive role to play in the presence of an optimal nonlinear income tax. Boadway and Pestieau (2003) show, however, that this result is not robust to various generalizations—such as unobserved differences in needs—and the empirical evidence is not supportive of such separability (Browning and Meghir, 1991). This implies that there is in principle some potential gain from differential commodity taxation, but in practice the information is rarely available to allow such fine-tuning to be made with any great confidence—beyond, perhaps, that usually achieved by a few key excises. The optimal tax literature thus continues to suggest, as observed by Kay (1989), that, practical considerations aside, there is little that commodity taxation can do which cannot be achieved equally well, or better, by an income tax and a few excises.

The case for employing a general commodity tax in addition to an income tax must thus largely turn on administrative practicality and enforceability. As Kay points out, the VAT has typically had some merit of simplicity in requiring much less judicial interpretation than has been involved in defining income for the purposes of income taxation. Even more fundamentally, VAT and other systems of indirect taxation rest on sources of information which are to some degree distinct from those used in administering income taxes. Spreading enforcement risk across multiple taxes may thus reduce the overall exposure of the fiscal system to revenue losses from evasion. The point is developed formally by Boadway, Marchand, and Pestieau (1994), who show (taking an extreme case for clarity) that when income tax can be evaded but commodity taxes cannot, welfare is improved by the introduction of a small uniform tax on consumption. Beyond this, moreover, information obtained in enforcing commodity taxes—on a trader's turnover, for example—may be helpful in enforcing the income tax (and vice versa). These arguments for deploying a range of instruments provide a powerful counter to the suggestion sometimes heard in the United States that the VAT should be seen as an alternative to the income tax: there may be very good practical reasons to have both.

The superiority often claimed for the VAT over the principal alternative form of indirect taxation, a single-stage Retail Sales Tax (RST), also rests on enforcement concerns. Applied at the same rate, a perfectly functioning VAT would collect the same revenue as a perfectly functioning RST (albeit with different timing), and have precisely the same incidence: implementation costs aside, the two are economically equivalent, as the stylized example in Box 2 illustrates.

Box 2. Equivalence of VAT and RST—When They Both Function Perfectly

The table below charts the application of a VAT and RST to a simplified chain of production, with one intermediate goods producer, selling all its \$10m output to a single producer of final consumption goods, which in turn sells all its \$15m output direct to the public. With a 20 percent rate of VAT, total revenues collected are \$3m, which is the same revenue as would be collected from a 20 percent RST rate applying only to final sales. The only difference is in the timing of tax payments, some of the VAT revenue being collected at an earlier stage in the chain of production, with a potential cash flow gain for the government.

	Intermediate Goods Producer	Final Goods Producer
Value of sales to final consumers	0	\$15m
Value of sales of intermediate goods	\$10m	0
Value of purchased inputs	0	\$10m
Output VAT on sales	\$2m	\$3m
Input VAT credit	0	\$2m
Net VAT due	\$2m	\$1m
RST on retail sales	0	\$3m

The differences between a VAT and an RST thus lie in the challenges for administration and compliance that they imply, and in the opportunities and incentives for evasion that they create. In administrative terms, the number of firms to be controlled under a VAT may be greater or less than under an RST: it will be greater in so far as all types of businesses, not just retailers, are brought into the tax, but lower to the extent that VAT systems typically exclude many small firms. The key practical distinction between the VAT and the RST is that revenue is protected under the former, but not under the latter, by levying tax at each stage of production: if the final seller is not taxed, all revenue is lost under an RST, but the fractional nature of the VAT means that tax would then be lost on the value added at that final stage (so long as VAT has been properly collected throughout the preceding production chain).

Some advocates of the VAT go further, suggesting that the VAT is ‘self-enforcing’ in the sense that each trader has an incentive to ensure that its suppliers have themselves properly paid VAT, in order that they themselves can claim an appropriate credit.⁶ There is an element

⁶ At the opposite extreme, it is also sometimes claimed that the ability to cross-check invoices—verifying that every credit claim is matched by some payment of output tax—can make the VAT especially abuse-proof. While there is again an element of truth in this—invoices do indeed provide an useful trail for VAT auditors—this too can be overstated. Such massive cross-checking has failed where it has been tried (notably in Korea in

(continued...)

of truth in this. Certainly businesses registered for VAT gain nothing (beyond some cash flow advantage) by purchasing inputs on an untaxed, undocumented basis, since they are in any event able to claim credit or refund for any VAT so paid. Moreover, there is an important sense in which the VAT is self-correcting, if not self-enforcing: if for some reason a supply to some registered trader escapes VAT, that missing VAT will be recovered at the next stage in the VAT charged by that trader on their own sales, since there will, in that case, be no credit to offset against their liability. For all these reasons, traders selling to other businesses have an incentive to register to charge the VAT even if their annual turnover is below the threshold at which VAT registration is mandatory, and indeed arrangements for such voluntary registration are a key part of any well-designed VAT.

But the strength of these intrinsic features of the VAT should not be over-stated. It remains the case that sellers of final goods to private individuals and businesses exempt from the VAT have the same incentive to sell without tax as they would under an RST (albeit muted to the extent that they would then also be unable to recover VAT on their own purchases). Moreover, while traders have an incentive to ensure that their suppliers provide them with invoices that the authorities will accept as establishing a right to refund or credit, they have no incentive—unless specific requirements to this end are imposed—to ensure that tax has actually been paid: for this reason, as Hemming and Kay (1981) stress, the notion that the VAT is self-enforcing is ultimately ‘illusory.’ The point will be of some importance below. Furthermore, the credit and refund mechanism of the VAT creates its own opportunities for fraud, as we now discuss.

A typology of VAT fraud and evasion

There are many ways in which the VAT can be evaded or fraudulently exploited. To derive a sense of the main risks, it is useful to distinguish between those that also arise under other forms of sales tax, the RST again being a focus of interest, and those reflecting distinctive features of the invoice-credit VAT.⁷

Frauds that can arise under both a VAT and an RST

- ***Under-reported sales.*** A trader may report only a proportion of sales, falsifying records and accounts to match, or may make some sales ‘off the books’ entirely.⁸

the late 1970s), and is likely to remain some way from feasibility for the foreseeable future, even with continuing dramatic advances in information technology.

⁷ The listing that follows is not intended to be complete: Tait (1988) lists sixteen types of VAT evasion, but there are others too: illegal transactions, for instance, are usually, in principle, subject to VAT. And there is always an incentive to find new ways of cheating.

⁸ Under both a VAT and an RST, traders may choose to support understatement of sales by also understating their purchases. While the gain from doing so is reduced under the VAT by the consequent inability to claim credit for VAT charged on inputs, a business that can make unrecorded purchases from a supplier not charging VAT may be able to operate a part of its business activity completely ‘off the books’ and without incurring unrecovered VAT. Indeed in this way one can envisage the emergence of whole production chains outside the

(continued...)

Under a VAT, they may or may not issue an invoice. If they do, the customer, if registered for VAT, may seek the corresponding credit. The ultimate success of the evasion then depends on the inability of the revenue authorities to discover that more VAT invoices have been issued but not declared by their issuer. For sales to final consumers or exempt businesses, however, there is no gain from issuing an invoice, since no credit would be available: common examples include personal services (hairdressing, home decoration, building contractors working for private customers, and so on) for which value added at the final stage is usually large relative to VATable inputs. Sometimes, of course, the customer is aware that the sale is made without VAT, and may share some of the gains (or believe that they do so). In many EU countries the customer in this situation is doing nothing illegal, though some have legislative provisions designed to discourage customers from making untaxed purchases. Italy, for example, requires customers to carry a receipt when leaving bars and restaurants.

- ***Failure to register.*** The commonest such cases under the VAT are relatively small businesses operating close to the level of turnover at which registration becomes compulsory, that fail to register, saving both the VAT for which they would be liable and VAT compliance costs. “Ghosts”—traders wholly unknown to the revenue authorities—may be able to evade income taxes as well as VAT. Once again, firms selling to final consumers (or to other unregistered businesses) are likely to predominate in this group.
- ***Misclassification of commodities.*** When traders have sales that are liable to tax—whether VAT or RST—at different rates, or some of which lie outside the scope of the tax (exempt items under the VAT), they may reduce their liability by exaggerating the proportion of sales in the lower-taxed categories.
- ***Omission of self-deliveries.*** Goods or services produced by the business and consumed by the proprietor or employees, in principle taxable, may not be declared. Generally, however, this is a relatively unimportant category in more developed economies where agriculture, for which this can be a real issue, is less significant. And a VAT with a relatively high registration threshold will exclude many of the smaller businesses for which this is likely to be more of a risk.
- ***Tax collected but not remitted.*** This may be possible either through false accounting (under-reported sales, as above), by engineering bankruptcy before tax is paid, or in other ways. More particularly, the ‘missing trader’ frauds discussed later typically involve registered businesses charging their customers VAT but disappearing before paying tax to the authorities. The same could also arise, of course, under an RST.

VAT, with a converse of the usual self-enforcing argument also applying: while an honest trader would prefer their supplier and customer to be honest, a dishonest one would prefer them to be dishonest. This risk of emergence of chains outside the VAT system can be a particular concern, as Tait (1988, p. 308) observes, in developing countries, where the chain of production between primary producer and final consumer is often short.

- ***Imported goods not brought into tax.*** If tax is not levied at the border, then there is a potential gain from purchasing imported goods bearing no tax and then reselling them in the home market. This is a risk under the RSTs of the United States, and indeed the constitutional obstacles to levying tax on out-of-state sales has made this a significant problem even internally. Most VATs, in contrast, bring imports immediately into tax and so are more robust against this type of fraud. As will be seen, however, the risk of such ‘acquisition fraud’ does arise under, and is a major weakness of, the current deferred payment system of the EU.

While there are, thus, many types of abuse that can arise under both a VAT and an RST, they seem likely to prove less extensive—all else equal—under the former than under the latter. And there are some types of fraud that are far more characteristic of an RST than of a VAT. For example, except for relatively limited instances of “reverse charging” (explained and discussed later), the VAT makes no distinction between sales by the nature of the purchaser,⁹ and so, unlike an RST, does not offer any opportunity for reducing tax payments by misclassifying sales to final consumers as business-to-business (B2B) transactions. Of particular interest here, however, are the opportunities for fraud that are distinctive to the VAT.

Frauds distinct to the VAT

At the heart of the VAT is the credit mechanism, with tax charged by a seller available to the buyer as a credit against their liability on their own sales and, if in excess of the output tax due, refunded to them. This creates opportunities for several types of fraud characteristic of the VAT.

- ***False claims for credit or refund.*** This is the most obvious way to exploit the credit mechanism: “A VAT invoice...” as Bird (1993) puts it, is “a check written on the government.” And such fraud can be extensive. Tait (1988, p.307), for example, reports that 44 percent of all VAT fraud found in an investigation in the Netherlands took the form of false claims for tax paid at previous stages, for example by presenting forged invoices for non-existent or exaggerated purchases. For established businesses liable to VAT on their own output there is some limit to this, since excessive credit claims would imply that the business was operating at an implausibly low margin. But there is much more scope for new businesses to exaggerate input claims without arousing suspicion, since large initial purchases of capital goods and other inputs would be expected at start-up, without being matched immediately by a corresponding level of sales.¹⁰

⁹ The ‘VIVAT’ discussed below, however, would rest on such a distinction, and so would be vulnerable to this type of fraud—though, as will become clear, this may be more than outweighed by a lesser vulnerability to more damaging forms.

¹⁰ For this reason, some countries do not refund VAT in these circumstances but instead either carry the credit forward until some output tax arises or exempt imports of this kind.

The universal zero-rating of exports means that fraudulent claims to have exported commodities are an intrinsic difficulty for the VAT: this, to give but one example, was the nature of the fraud in Sri Lanka referred to above. Such fraud may involve non-existent commodities, or genuine commodities that have, in fact, been sold on the domestic market (“diversion fraud”). In either case, it poses a challenge in implementing the VAT that all tax administrations find trying: between the desire to pay refunds promptly, in order to avoid converting the VAT into an export tax, and the need to protect revenues.

A false VAT credit claim for \$1 costs the government the same, of course, as a fraudulent refund of \$1. For the criminally inclined who would otherwise be outside the tax system, however, it is only the latter that will be tempting. And, thus, it is that the problems created by VAT refunds commonly receive particular attention.

- ***Credit claimed for VAT on purchases that are not creditable.*** This arises in two main forms. First, when traders supply a variety of outputs, some subject to VAT and others exempt, they have an incentive to allocate inputs to production of the taxed items (in respect of which input tax credit is available) rather than the exempt (for which it is not). While rules are commonly adopted for such apportionment, the borderline inevitably proves contentious—and indeed there may well be no conceptually correct allocation. Second, items bought for private consumption may be misrepresented as business inputs, allowing the VAT to be recovered (and income tax liability reduced).
- ***Bogus traders.*** Companies may be set up solely to generate invoices that allow recovery of VAT. Such “invoice mills” exploit the practical impossibility of cross-checking every invoice against evidence that earlier tax has been paid.

What matters, of course, in comparing the VAT and its alternatives is not the number of possibilities for evasion, but their quantitative significance. This is likely to be affected by the number of taxpayers in a position to exploit particular possibilities for abuse, the risks of detection and enforcement action that they face—and, not least, the way in which the taxes are designed.

Implications for VAT design

It is clear from the account above that the design of a VAT can significantly impact its exposure to evasion. Key design aspects from this perspective include:

- ***Rate differentiation.*** Multiple rates create scope for misclassification fraud of the kind mentioned above. Moreover, where rate differentials are sufficiently large they can also give rise to refund entitlements for some traders (those using inputs taxed at a high rate to produce outputs subject to a low rate of VAT), which, in turn, creates opportunities for fraudulent abuse. Although a reduced rate of VAT is the most

elegant way¹¹ of taking account of distributional considerations in the design of a VAT (since it avoids the production inefficiency associated with the taxation of inputs implied by exemption), it has the disadvantage of extending the scope of the refund problem beyond exports. As can be seen from Table 1 of Keen and Lockwood (2006a), VATs in the EU are indeed characterized by considerable rate differentiation. This largely reflects their age, and the difficulty of removing privileges once given: a notable feature of most newer VATs is that they are commonly characterized by a single non-zero rate (International Tax Dialogue, 2004).

- ***The level of VAT rates.*** Several frauds become more attractive, of course, the higher is the rate of VAT. High VAT rates are likely, in particular, to encourage informality: although they raise the unrecovered tax that informal operators will bear on their inputs they also increase the output price they can charge while still undercutting formal operators, with the latter effect outweighing the former so long as their own value added is positive. Such considerations naturally point to a lower VAT rates than would otherwise be the case.¹²
- ***Exemptions.*** Though less elegant than a reduced rate, exemption may be a more practicable way of dealing with distributional objectives since it obviates any need for refunds. Exemption is often also used to circumvent a variety of technical difficulties, including that of properly taxing financial intermediation, though recent years have seen movements towards some input tax recovery in this sector.¹³ VAT exemption is not without its difficulties, however, in both the production inefficiency it potentially induces and the problems that arise in allocating input VAT for producers of taxed and both goods.
- ***Registration thresholds.*** The number of firms that have to be handled by the VAT administration can be sharply reduced by setting a high turnover threshold. The VAT registration threshold varies widely even within the EU, from around UKL 61,000 (about \$115,000) in the United Kingdom to zero elsewhere. Outside the EU, too, some very high thresholds can be found: around \$100,000 in Lebanon, for example, and \$500,000 in Singapore. The revenue lost by setting a high threshold may be small compared to the saving of administration costs to the authorities and compliance costs to the taxpayer, because the potential tax base is commonly very strongly concentrated in the largest companies. Moreover, firms not registered for VAT nevertheless face a non-zero effective rate of tax, because they cannot reclaim the VAT paid on inputs. Also to be factored into setting the threshold, however, is the potential distortion that any threshold creates in the competition between those above

¹¹ Developed economies, however, are likely to have better-suited instruments to hand, so that sales tax differentiation will rarely be the best way of pursuing distributional objectives: see Ebrill and others 2001, Chapter 7.

¹² For a formal treatment of this, see Keen (2006).

¹³ On the issues and current practice in taxing financial services under the VAT, see for instance Zee (2006).

and below it; this will be eased by the possibility of voluntary registration, but it is by no means yet resolved whether the appropriate response to this distortion is to set the threshold lower than would otherwise be appropriate.¹⁴

- ***Simplified and lump-sum schemes for small traders.*** The revenue loss from nonregistered firms can be reduced by various special tax regimes for smaller firms, such as those that simply tax smaller firms on their gross turnover. Such schemes may offer both firms and the authorities substantial savings, while reducing both evasion opportunities and competitive distortions.
- ***Timing of payments and refunds.*** The scope for some VAT frauds is affected by the speed with which VAT refunds are paid, compared with VAT collections. Frauds that involve false refund claims by firms which subsequently disappear will be more tempting the more rapidly refunds are paid, both because this gives the authorities less time to detect the fraud and because a longer time lag in refunds compared with collections means that firms may, for a period, be more substantial net tax creditors to the tax authorities. Again, however, there is a delicate balance to be struck: it may be that the difficulties currently experienced in the United Kingdom, for example, in part reflect the authorities' strong commitment to prompt payment of refunds; in many developing countries, on the other hand, many traders find refunds extremely difficult to obtain (and in some cases they have a well-established price in terms of bribes to be paid).

Taken together, the general structural characteristics of the VAT—argued above to make it less exposed to evasion than the RST—and the use of these various design opportunities for reducing its vulnerability to evasion have permitted many European countries to levy VAT at high rates with a reasonable degree of success. Recent concerns in the EU, however, have certainly dented any sense of complacency, and it is to these that we next turn.

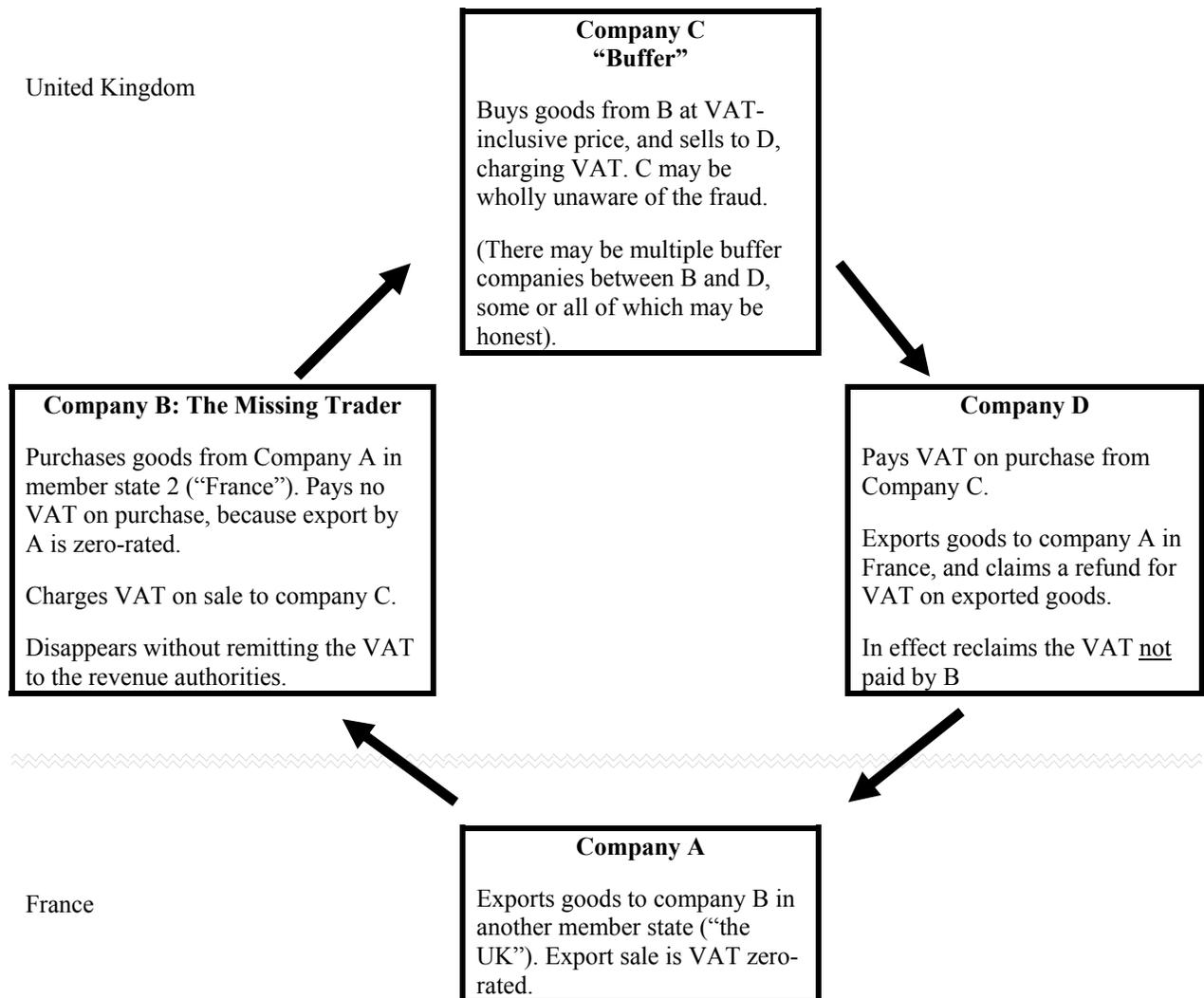
III. INTERNATIONAL ASPECTS OF VAT FRAUD

As discussed above, the frauds most distinctive to the VAT exploit are the crediting of tax paid on purchased inputs against the tax due on sales. And it is the zero-rating of exports that has proved the feature most vulnerable to fraud, or at least which has proved especially attractive to deliberate criminal attack. The difficulty with zero-rating exports is that it not only breaks the VAT chain but does so “at a particularly vulnerable spot: the interface of domestic and foreign tax administrations” (Ainsworth (2006), p.445). This creates a number of opportunities for fraud, as outlined above. But the best known form of deliberate attack on the VAT system is the “carousel” fraud that has attracted such attention in the EU over the last few years.

¹⁴ Formal analyses of the optimal choice of threshold are provided by Keen and Mintz (2004), Dharmapala and Slemrod (2006), and Zee (2005).

“Carousel fraud”—a special case of what is termed in the United Kingdom “missing trader intra-community fraud” (MTIC)—exploits the zero-rating of exports combined with the “deferred payment” mechanism for collecting VAT on imported goods. Under the latter, adopted in the EU with the removal of fiscal frontier formalities in 1992,¹⁵ VAT on imports from another member state is collected not at the border but at the time of the next periodic return. In the simplest case, illustrated in Figure 1, carousel fraud works as follows:

Figure 1. The Basic Carousel Fraud



¹⁵ The EU abolished fiscal frontier formalities between member states in 1992, as part of a wide-ranging package of measures designed to promote economic integration between member states—the “Single Market” program. Member states are not permitted to operate frontier formalities, including fiscal procedures, which could impede the free flow of goods trade between them. As a result, verification of the right to zero-rate goods exported from one member state to another can no longer make any use of border procedures, and must be based on audit of documentary evidence from the seller and/or buyer. Member states are, however, still able to use frontier formalities to verify the export of goods outside the EU.

Step 1: Company A, located (and registered for VAT purposes) in, say, France, exports some good to company B located in, say, the United Kingdom (and registered there). As a consequence of export zero-rating, A receives a full refund of any input VAT paid. The VAT-free unit price at which the good is sold is, suppose, €100.

Step 2: Under the deferred payment scheme, company B is not required to pay VAT at the time of import, but rather to account for this in its next periodic return. Suppose that the applicable rate of VAT in the United Kingdom is 15 percent, and (for simplicity) that company B adds no mark up on its sale. Then it will sell on to another United Kingdom company, C, charging it €115, with €15 of this shown as VAT.¹⁶ The invoice issued to C entitles it to a €15 credit against its own output VAT. But now comes the catch: firm B does not remit to the tax authorities the €15 it has charged and invoiced—instead, it goes missing without paying any VAT.

Step 3: The invoice issued by B continues its way down the production chain, through firm C and perhaps other ‘buffer’ companies—which may be unaware of the fraud being perpetrated—until it arrives at firm D, which now exports the good out of the United Kingdom. In doing so, it obtains a VAT refund for an amount that includes the VAT which firm B, in fact, did not pay. If (to keep the numbers simple) these buffer firms also add no value, so that D purchases at a VAT-inclusive price of €115, then the amount of the refund is €15, corresponding to the VAT that the importing firm B invoiced but failed to remit.

Step 4: In a pure carousel fraud, the export by D takes the goods back to France (perhaps via third countries). The whole process starts again, with the goods—usually high value items relatively cheap to transport, such as mobile phones or computer chips—going round in a lucrative circle.

The scheme just described is, by the high standards of many criminal operators, extremely simple: in practice, the fraud can be immensely obscured by using complex buffer operations and country routings (the export by firm D, for example, could be to a non-EU country, since that also gives rise to refund), and there may be a ‘price drop’ at the import stage—B charging a VAT-exclusive price of less than €100—both to make participation by C lucrative and to ensure that the exporting firm A does not incur suspicion by showing a perennial commercial loss. It does though bring out three key generic features of the carousel fraud:

- This is, in effect, a “double-dip” fraud. First, firm B steals a competitive advantage relative to legitimate traders—to the tune of €15 in the example above—by keeping for itself what an honest business would remit as tax. This stage of the operation—the acquisition fraud—would be remunerative even if there were no subsequent export. The second element is the cash refund at export, worth, in the example above, another €15.

¹⁶ The VAT due is calculated as the sum of that due on importation but deferred, €15, plus the excess of any output VAT due over that import tax—which in this case is zero.

- Viewed from outside, it may be far from obvious which traders—other than the initial importer, who has simply disappeared—are party to the fraud. In the example above, those in between the import and export stage could be perfectly innocent (though in practice they may have been offered suspiciously attractive prices for their purchases). And, of course, in a simple acquisition fraud both the initial exporter A and the final exporter D could also be innocent.
- The key player in the fraud—the importer who is going to go missing—needs to be registered for the VAT. This is because EU rules allow zero-rating of exports to another member state only on production of a VAT registration number for the purported importer: giving a fraudulent number would expose A to risk of detection.

But the final generic feature is perhaps also the most important: carousel fraud is very clever. In considering how to deal with it, the resourcefulness and creativity of the authorities' adversaries should not be underestimated.

IV. THE EXTENT OF VAT NONCOMPLIANCE AND FRAUD

Even allowing for the intrinsic difficulty of measuring fraud and evasion, there is strikingly little hard evidence—publicly available, at least—on the extent of noncompliance (including outright fraud) under the VAT. Numbers are quite commonly mentioned: the European Commission (2004), as noted above, reports that losses from fraud have amounted to 10 percent of net VAT receipts in some member states, and the German Federal Audit Court has put losses there in the region of 11 percent of VAT revenue in 2003.¹⁷ But the precise basis of such figures is often unclear (perhaps for good reason, in that it may reveal information on how the fraud is being combated). This section reports on some of the few estimates whose basis is clear, focusing too—given the wider relevance of and current interest in the issue there—on the evidence emerging in the EU.¹⁸ It then asks, very briefly, whether noncompliance under the VAT is notably more or less than under other taxes.

The United Kingdom: official estimates and methodology

Since 1992, Her Majesty's Revenue and Customs (HMRC) in the United Kingdom has published an annual estimate of the VAT 'gap' in the United Kingdom. This compares actual VAT receipts with an estimate, largely from national accounts data on household spending,

¹⁷ *Tax Notes International*, August 8, 2005, p.489.

¹⁸ Being less immediately relevant for the United States, we do not consider here estimated VAT gaps in lower-income economies. These tend to be much higher than those reported below for most EU members, as one might expect. Brondolo and Silvani (1993), for example, usefully report compliance estimates, calculated on a consistent basis (by broadly the method described above), for 20 countries in the early 1990s: average noncompliance for those that are not now OECD members (almost all in Latin America) was about 35 percent, and above 50 percent in some.

of the hypothetical VAT revenue that would be obtained with full compliance: the “theoretical total VAT liability” (VTTL) in the United Kingdom’s jargon.

The broad methodology used is straightforward, although in practice a number of steps involve significant judgement and imprecision. It is a “top down” approach, starting with macroeconomic data on consumer spending, broken down according to the pattern of VAT rates. Applying the relevant tax rates (and taking account too of exemptions) to each category of final spending gives an initial estimate of VAT revenue. Four main adjustments are then made to compute the VTTL. First, amounts are deducted in each taxed spending category to reflect the exemption of sales by small businesses genuinely below the VAT registration threshold: these firms also cannot recover input VAT, so face an effective rate of VAT likely to be in the order of 3–6 percent (depending on sector). Second, an adjustment is made to reflect the effective rate of VAT on supplies of explicitly exempted items, such as the financial sector, health and education services, and clubs. Third, timing adjustments are needed to reflect the average lag of approximately one quarter between a transaction and the corresponding VAT receipt. Fourth, adjustments are made to reflect miscellaneous legitimate recovery of input VAT by certain types of business, mainly those supplying government departments and broadcast companies (BBC, ITV).

Resultant estimates of the VTTL, the revenue shortfall and the percentage revenue gap from the 2005 exercise are shown in Table 1. In the two most recent years, the tax gap has been some 13.5 percent of the VTTL, an annual revenue loss of approximately £11 billion.

Table 1. HMRC’s “Top Down” Estimate of the United Kingdom VAT Gap, Tax Years 2000–01 to 2004–05

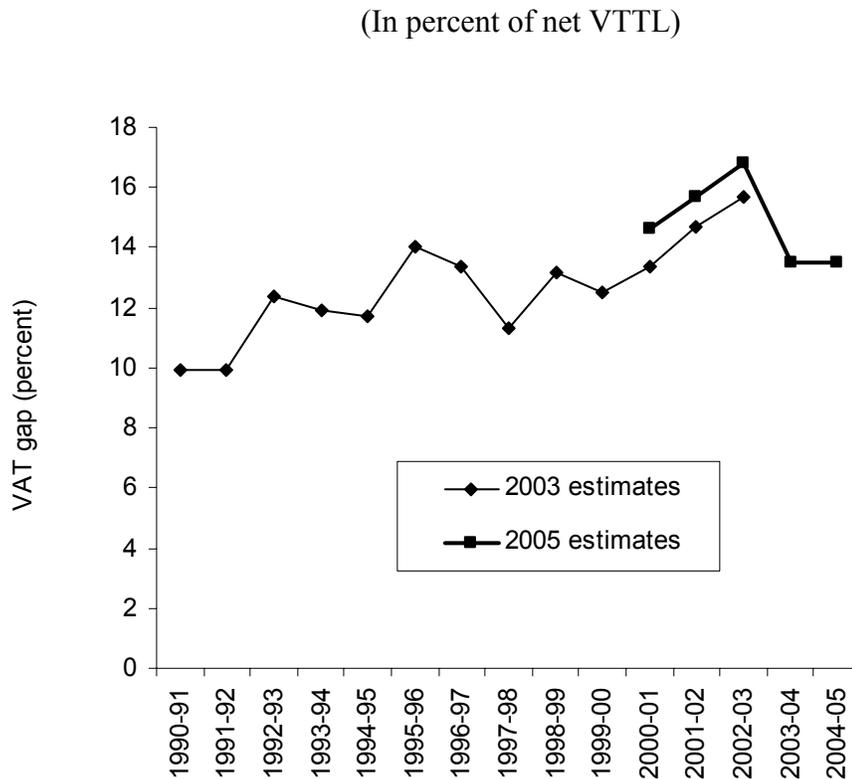
	Net VTTL (£ bn)	Net VAT receipts (£ bn)	Revenue shortfall (£ bn)	Tax gap in percent of net VTTL
2000–01	68.5	58.5	10.0	14.6
2001–02	72.4	61.0	11.3	15.7
2002–03	76.5	63.6	12.8	16.8
2003–04	79.9	69.1	10.8	13.5
2004–05	84.0	72.7	11.3	13.5

Source: HM Revenue and Customs (2005), **Table 2.1**

Estimates have been made on a similar basis for each year since 1990, but a fully consistent times series is not available, because of adjustments and revisions which have been made at various times. HMCR (2003) presents a time series for the percentage tax gap for each year from 1990–01 to 2002–03 (which is not wholly consistent with the figures in Table 1,

because some significant upward revisions have since been made for recent years). These estimates, illustrated in Figure 2, point to an upward trend in the VAT gap, from just under 10 percent in the early 1990s to a peak of around 16–17 percent in 2002–03. This would be consistent with an increase in opportunities for evasion following the abolition of EU internal border controls at the start of 1993 (discussed above), while the fall in evasion over the last few years—perhaps now reversed—may reflect the authorities’ conscious targeting of VAT evasion.

Figure 2. The VAT Gap in the United Kingdom, 1990–01 to 2004–05



Sources: 2003 estimates: HM Customs and Excise (2003), [Table 3.1](#); 2005 estimates: HM Revenue and Customs (2005), [Table 2.1](#).

Given the imprecision of, and frequent adjustments in, the underlying national accounts data, and the limited information on which to base some of the above adjustments, these estimates clearly have quite a wide margin of error. And indeed HMRC see them as more reliable in indicating changes in VAT evasion than in its level. Nevertheless, some confirmation of the level estimates has been obtained from a separate “bottom-up” exercise, in which operational data were used to assess the likely order of magnitude of VAT evasion in a number of major categories. In addition to providing an independent check on the plausibility of the “top-down” approach to estimating the VAT gap using national accounts data, the bottom-up estimates also provide better diagnostic information about the quantitative significance of different types of VAT fraud.

Table 2 reports these bottom-up estimates for the tax year 2001–02, under four separate headings: missing trader fraud, “artificial” tax avoidance arrangements instituted solely to reduce VAT liability, nonregistration by firms with turnover above the VAT registration threshold, and general noncompliance by VAT-registered firms (for example by understating sales or overstating purchases, either intentionally or by mistake). These estimates are based on a range of information sources, including operational data, audit results, and audit. The methodology is not fully disclosed, because of concerns about the operational sensitivity of some information sources. Reflecting their imprecision, the estimates are reported as a range of values.

Table 2. Bottom-Up Estimates of VAT Revenue Losses, United Kingdom, Tax Year 2001–02

Type of Noncompliance	Estimated Revenue Loss (£ bn)	Estimated Revenue Loss (In percent of full-compliance VAT revenues)
Missing trader (MTIC) fraud	1.77–2.75	2.5–3.9
“Artificial” tax avoidance	2.5–3.0	3.6–4.3
Nonregistration for VAT	0.4–0.5	0.6–0.7
General noncompliance by VAT-registered firms	2.5–4.0	3.6–5.7
Total	7.17–10.25	10.2–14.6

Source: National Audit Office (2004), page 11.

For comparability with those shown above, “the bottom up” estimates are shown in Table 2 as a percentage of VTTL (assumed to be £70 bn). Overall, this approach indicated VAT revenue losses of about £7–10 billion, equivalent to a VAT gap of 10.2–14.6 percent of VTTL. The upper end of this range is broadly consistent with the 2001–02 “top-down” estimate of £11.3 billion, or 15.7 percent of VTTL.

Within this total, missing trader fraud accounted for revenue losses of 2.5–3.9 percent of the VTTL (or £1.8–2.75 billion) in 2001–02. The basis of the estimate of missing trader fraud is only partly disclosed, but the upper point of the range has been estimated from discrepancies between statistical data—which are based on VAT reports—on sales to the United Kingdom member states and purchases from other member states declared in the United Kingdom. (The point here is that the exports associated with carousel fraud—which are declared (by firm D in the example above) in order to benefit from the refund of VAT—will be reflected in the base data used to estimate United Kingdom trade statistics, while the corresponding imports by missing traders (firm B in the example) will not. Thus the United Kingdom’s trade deficit will be understated). More recent estimates of MTIC fraud for 2000–01 to 2004–05 (HM Revenue and Customs, 2005), suggest that the revenue loss from such fraud

fell quite significantly to some £1.12–£1.9 billion in 2004–05, equivalent to a VAT “gap” from this source of 1.3–2.3 percent.

Further, and more disquieting, indications of developments in missing trader fraud in the United Kingdom are provided by recent adjustments made by the Office of National Statistics (ONS) to published trade data. These show a sharp growth in fraudulent trade, from £2.7 billion in 2004, to £11 billion in 2005, and £6.5 billion in the first quarter of 2006 alone (Bank of England, 2006, p 23). If the latter figure were to be maintained through the year, it would imply fraudulent exports of £26 billion, and, applying the U.K. standard VAT rate of 17.5 percent, associated revenue losses of some £4.5 billion, about 5 percent of VTTL—which is about twice the upper bound of HMRC’s estimate of the revenue lost through missing trader fraud in 2004–05.

National accounts-based estimates for selected EU countries

Using a “top-down” approach broadly similar to that described above, Gebauer and Parsche (2003) estimate VAT evasion rates for 10 EU countries. Table 3 reports the evasion rates for each of the 10, averaged over the three years for which estimates have in each case been made.

Table 3. Estimates of the VAT Gap for Ten EU Countries

(Three-year averages)

	VAT Gap (Evasion) as Percentage of Hypothetical Revenue	Estimate Based on Years
Netherlands	2.4	1994–96
Great Britain	3.8	1991–93
Denmark	4.2	1994–96
Germany	4.8	1994–96
France	8.8	1991–93
Portugal	14.2	1994–96
Belgium	19.3	1994–96
Greece	20.2	1994–96
Spain	22.6	1994–96
Italy	34.5	1991–93

Source: Calculated from data in Gebauer and Parsche (2003), Table 1.

There are clearly substantial cross-country differences. The Netherlands, Great Britain, Denmark and Germany have average evasion rates of under 5 percent, while evasion rates

exceed 15 percent in Belgium, Greece, Spain and Italy. In Italy, most notably, the estimated evasion rate implies that for every billion Lire collected in VAT, a further half-billion Lire in tax was evaded; this is broadly consistent with Pedone's (1981) earlier estimate that 40 percent of potential VAT revenue in Italy was lost through evasion.

Across the three years for which the study reports estimates, the VAT gap was fairly stable in most countries, but was found to have increased sharply in both Germany and the United Kingdom. In Germany, the VAT evasion rate rose from 1.6 percent to 7.5 percent over the three years 1994–96. Figures for Germany for later years reported by Gebauer and Parsche (2003, Table 2), based on data that are not wholly comparable with the earlier period, suggest that the growth in VAT evasion slowed in the late 1990s, but that the VAT gap then rose sharply to 9.5 percent in 2001—broadly consistent with the later estimate for 2003 cited above. More recently, the same methodology suggests a fall in VAT evasion in Germany in 2005, in sharp contrast with experience in the United Kingdom.

Strikingly, Gebauer and Parsche (2003) put the VAT gap in the United Kingdom at only 0.4 percent in 1991, rising to 4.4 percent in 1992 and 6.5 percent in 1993. These figures are much lower than the official United Kingdom estimates for these three years reported above, which are roughly three times higher on average across the three years (and more than 20 times higher for 1991). And while the official United Kingdom estimates also show a rising trend over a number of years, they do not show anything like the same absolute increase in VAT evasion over these three years, and nor do they show the same pattern of increase. The reasons for these differences cannot be tracked from the published sources, but seem likely to reflect different judgements about the various adjustments made in moving from macroeconomic data on sales to the theoretical full-compliance level of VAT revenues. In interpreting the figures for other countries in Table 3, the possibility of a similar margin of error needs to be borne in mind. Estimating the VAT gap—like other attempts to measure economic activity deliberately concealed by the perpetrators—is inevitably an imprecise art.

Econometric analyses

One would like to know, of course, not only the extent of VAT noncompliance but also something of its determinants. Is there any sign, for example, that it is in fact significantly affected by some of the design aspects raised above?

The difficulty of measuring VAT noncompliance, evident from the discussion above, has impeded serious empirical work on these questions. As far as we are aware, no researcher has assembled more than a small cross-section of observations of VAT compliance. Agha and Haughton (1996) construct and analyze such estimates for 17 OECD members, in 1987. They conclude that (consistent with the presumption noted above) a higher VAT rate is associated with lower compliance, with an effect that is both large and statistically significant: a 1 percentage point increase in the VAT rate (from the sample average of 15.8 percent) would reduce the compliance rate by 2.7 percentage points. They also find that the number of VAT rates has a significant and large effect on compliance (each additional rate reducing compliance by, on average, 7 percentage points), and that compliance increases the longer the VAT has been in operation, and the smaller the country in terms of population. Other

factors that might influence compliance, including the VAT base as a proportion of GDP, the severity of penalties for late payment, and the proportion of the population registered to pay VAT, had no significant impact on compliance. Given the small sample size, however, and potential heterogeneity bias, these conclusions need to be interpreted with great caution. A similar exercise had earlier been undertaken by Brondolo and Silvani (1993), although on a much more diverse group of 20 countries. They find little evidence of significant design effects on consumption.

A few papers explore the determinants not of some direct measure of compliance with the VAT but rather of some measure of its revenue productivity. Recent work of this kind (Ebrill and others, 2001), Aizenman and Jinjirak (2005)) focuses on “C-efficiency,” defined as the ratio of VAT revenue to aggregate consumption, divided by the standard rate of VAT: under a uniform single rate VAT, perfectly enforced, C-efficiency would be unity. The advantage of analyzing C-efficiency is that it is relatively easily measured, giving the possibility of a reasonably large data set. The disadvantage for present purposes, however, is that C-efficiency depends not only on the level of compliance but also on the VAT rate structure: leaving aside exemptions, C-efficiency can be shown to be the product of a measure of aggregate compliance and the ratio of the average VAT rate (weighted by the shares of consumption taxed at each rate) to the standard rate. Thus, great care is needed (as both sets of authors recognize) in interpreting the coefficients on explanatory variables as indicating an effect on compliance. Both find, for example, that C-efficiency falls with the share of the agricultural sector in GDP—but it is unclear whether this indicates that compliance tends to be low in what is a classic “hard-to-tax” sector or, alternatively, because agricultural products are commonly subject to a low rate of VAT (or exempted). Some suggestive findings do nevertheless emerge: for instance, the significant positive impact of openness found by both sets of authors, at least for lower-income countries, is consistent with the widespread view that collection of the tax on imports can be key to implementation of the wider VAT.

Similar qualifications apply to an alternative approach which takes as dependent variable “VAT productivity,” this being VAT revenue divided by the product of the standard rate and GDP. Indeed, in this case, any effect from the standard rate, for instance, might also arise from induced movements in aggregate consumption (a general increase in the rate of VAT having incentive effects, as noted earlier, akin to those of an increase in the rate of labor taxation). While such analyses may be informative in understanding the revenue yield of the VAT—leading Matthews and Lloyd-Williams (2000) and Matthews (2003) to suggest a revenue-maximizing rate for EU countries of around 18–20 percent—they conflate compliance effects with other consequences of rate changes. All that can be concluded from these efficiency-based studies is that many of their findings¹⁹—such as negative impact of the standard rate, and from the range of rates applied—are broadly consistent with, but do not prove, many of the presumptions of practitioners.

¹⁹ Though not all: Ebrill and others (2001) find that an increase in the range between highest and lowest VAT rates *increases* C-efficiency (noting that this is not necessarily surprising, since uniform taxation may not be revenue maximizing).

Comparison with other taxes

The vulnerability of a tax to fraud and evasion is not, of course, the only consideration in determining whether, and to what extent, it ought to be deployed. A tax may be hard to evade, for instance, yet cause significant inefficiency or unfairness. Nevertheless, it is natural to ask whether the VAT seems to be more or less vulnerable to abuse than are alternatives.

In terms of commodity taxation, the natural comparator in the United States context is the RST. Hard evidence on the extent of noncompliance with the state sales taxes is surprisingly sparse, but the review in Fox and Murray (2004) tends to confirm a conventional wisdom that it is quite modest (at about 1.3 percent in Washington state, for example). Evasion of the use tax on out-of-state purchases—somewhat analogous to losses on intra-community supplies within the EU—is, of course, much greater (around 28 percent in Washington state). But the apparent modesty of the sales tax gap reflects the generally low level of rates (though that in Washington, at 6.5 percent, was by U.S. standards relatively high). And indeed another folk wisdom is that the RST becomes unworkable at rates of more than 10–12 percent—certainly no country that we are aware of currently levies an RST at such a level.

For the individual income tax, reasonably precise estimates for the United States put the tax gap at 13.7 percent (Slemrod, 2006). The corresponding figure for the United Kingdom is likely to be somewhat lower, given the lesser importance there of self-employment income and the greater use of withholding. Noncompliance in the United Kingdom may, thus, now be greater under the VAT than under the personal income tax—but it is not of a completely different order of magnitude.

What of other schemes which, like the VAT, involve payments to taxpayers? The leading example in the United States is the Earned Income Tax Credit, for which the most recent study of noncompliance (on returns filed in 2000) found that overpayments amounted to about 27–32 percent of the total paid (Internal Revenue Service, 2002). Much of this, it seems, represented genuine errors rather than fraud. In the United Kingdom, excess payments under the Working Families Tax Credit have been reported as running at around 10 percent of the total,²⁰ in this case it seems in part reflecting errors by the administration. While these figures are not directly comparable to those for the VAT gap above, they serve as a reminder that the VAT is not the only part of the tax system that may struggle to avoid extensive overpayment.

V. ENFORCEMENT STRATEGIES

The implementation of a VAT involves the same core elements as does any other self-assessed tax: the identification and registration of those required (or choosing) to pay the tax; collection and processing of amounts spontaneously remitted with periodic returns; audit to

²⁰ As reported by the BBC, July 11, 2006: <http://news.bbc.co.uk/2/hi/business/5168452.html>.

ensure accuracy of returns; and enforcement action on delinquent payers. These essential elements are the same, for example, as under the income tax, so are not rehearsed here.

Instead we focus on the distinctive considerations and operational challenges that arise under the VAT. These stem from two related sources. The first is the key role played by the VAT invoice in establishing both the tax due from the seller and the credit available to the purchaser, discussed at length above. The second is that taxpayers' liabilities can and often will be negative. In particular, zero-rating of exports unavoidably means that many taxpayers will systematically be due a VAT refund, and any domestic zero-rating will amplify this. While the likely importance of refunds depends on both the design of the VAT and structural features of the economy (being greater, in particular, the higher is the ratio of exports to GDP),²¹ they are commonly very sizable: Harrison and Krellove (2005) report that, for a sample of EU countries, refunds average a little under 30 percent of gross VAT collections—and around 50 percent in Canada and Sweden (reflecting their extensive zero-rating). As seen above, refunds open up particular revenue risks, placing the tax authorities in an intrinsically awkward situation: too lax an attitude towards the payment of refunds invites fraud, too miserly an attitude transforms the VAT into exactly what it is not supposed to be—a tax on production, and on exports in particular. Striking an appropriate balance between these two has proved one of the most difficult aspects of VAT administration. Countries have developed a range of devices to this end, which are thoroughly reviewed in Harrison and Krellove (2005). Among the least attractive is the zero-rating of “indirect exporters”—those supplying exporters (which can simply push the problem one step back down the production chain, and spread it to a larger number of taxpayers). Among the more promising are “gold card schemes,” under which those with good compliance records are assured prompt payment of refunds (though this can disadvantage new entrants who have not had time to build up a track record of compliance, and audit remains necessary to protect against strategic good behavior preparatory to substantial fraud). No method, however, is perfect.

Many of the core difficulties in dealing with refunds, and in VAT implementation more generally, are writ large by the carousel schemes described above, since these play off the zero-rating of exporters and the disconnect between the issuance of VAT invoices and the actual payment of tax. Schemes intended to deal with such fraud are thus of particular interest. They may also be of particular relevance to the United States, since if a VAT were to be run as—at least in part—a state-level tax, precisely the same issues would arise in relation to inter-state exports as now arise in the EU, since the absence of border controls would then create difficulties similar to those associated with deferred payment in the EU.

Broadly two types of solutions to current difficulties in the EU have been proposed. Some are essentially administrative, in the sense that they retain the zero-rating of intra-community supplies; the others are more fundamental, removing export zero-rating within the Union altogether.

²¹ Ebrill and others (2001) derive a simple formula relating the appropriate level of VAT refunds to such structural characteristics.

The first level of administrative response, quite naturally, is to do the usual things better, and if need be adapt the legal framework. One key focus of attention is registration, which is a prerequisite for these frauds: defensive measures might involve, for example, tighter checks on would-be registrants (for example, with an on-site visit and background checks), and requiring guarantees in dubious cases. Other administrative measures include adopting or strengthening joint and several liability rules by which traders can be held responsible for fraud elsewhere in the chain that they might reasonably have been expected to be aware of; and establishing better and quicker information exchange between national tax authorities (so that the country of import can promptly become aware that exports to it that have been reported in another member state have not shown up in its own VAT system).²²

More radical administrative proposals include:

- The use of ‘*reverse charging*’ for business-to-business (B2B) transactions, placing VAT liability on the buyer rather than the seller (with the amount thus paid fully creditable against the purchaser’s liability on any subsequent sales).²³ Imports, in effect, are conventionally taxed on this basis, with liability placed on the buyer. Extended throughout the production chain, reverse charging would frustrate the fraud illustrated in Figure 1, for instance, by passing to firm C the VAT liability on the sale by the soon-to-be-missing firm B: even if B vanishes without paying any VAT, the appropriate amount will be collected on C’s purchases. The United Kingdom has recently adopted reverse charging for mobile phones, computer chips, and other particular goods that have proved popular instruments for carousel fraud. More radically still, Austria and Germany have proposed allowing reverse charging for *all* B2B transactions above a certain (modest) size.²⁴

There are, however, a number of difficulties with reverse charging. By eliminating part of traders’ output tax liability, it may increase refund claims, with consequent control difficulties and risk of fraud. If applied only in respect of particular commodities, it may simply displace current difficulties onto other items. That dynamic is thus likely to create pressure to extend the scope of reverse charging—which, in the limit, would, in effect, convert the VAT into an RST, with the attendant difficulties that creates: in particular, with tax suspended on B2B transactions, if the final seller for some reason fails to remit the VAT due then no tax is collected. If

²² The EU maintains a VAT Information Exchange System (VIES) that enables exporters to verify that their customer in another member state is registered for VAT (so entitling them to zero-rating), and can be used to track transactions (checking that exports reported in one member state as going to another duly turn up as imports in the latter). But the lags involved in the latter function are currently such that the system is seen as more useful for ex post assessment of the extent of fraud than its ex ante prevention.

²³ The Rationalization Directive (adopted July 24, 2006) allows member states to reverse charge in identified sectors and for certain types of transactions, and indeed they do so quite extensively.

²⁴ €5,000 for goods, €10,000 for services. Such limits create further distortions in combining or splitting transactions.

applied wholesale, reverse charging would mean, to a large degree, the death of the VAT. For these reasons, the Austro-German proposal has been widely and strongly resisted.

- Similar considerations to those motivating reverse charging proposals might point to considering also the use of “*reverse withholding*” in vulnerable sectors. Such schemes—quite widespread in Latin America and Africa, though not in Europe—have some similarity with reverse charging, in that the purchaser is again required to remit payment of VAT in respect of its purchases—but the seller remains liable for output VAT, with a credit given for the amount withheld by the purchaser (which may be at lower than the normal VAT rate).

Reverse withholding can protect revenue more firmly than does reverse charging: the effect if firm B goes missing is the same as under reverse charging, so long as withholding is at the full rate, but not all revenue is lost if the final seller escapes tax. It does though potentially increase the need to pay refunds, since sellers in effect take credit not only for input VAT but also the VAT withheld by their customers. Moreover, the need to ensure that the seller is properly credited (or refunded) for the VAT withheld by its customers places additional demands on the administration: and to the extent that implementation of this is imperfect the effect is to transform the VAT, in part, into a tax on production.

- Adoption of a system of “*VAT accounts*”—as proposed for example by Sinn, Gebauer, and Parsche (2004), and as implemented in Bulgaria—under which traders would be required to open a distinct bank account into which they would transfer the amount of VAT charged to their customers. In the event of suspicious refund claims, the authorities could then simply check whether the amount claimed had in fact been paid. Thus the carousel frauds set out above, for example, would be stymied by the denial of credit to company C on the grounds that firm B had not paid the corresponding amount into the required account.

One evident disadvantage of this approach is that it imposes additional compliance costs on taxpayers, not only in the requirement to open an additional account but also, and probably more substantially, in the interest foregone through earlier payment of tax. VAT would become due once the sale is complete rather than with the periodic return. (The implied movement of the VAT from an accrual to a cash basis also raises difficulties in itself.) Moreover, it is not clear why it should be easier for the authorities to verify payments into a bank account other than their own. Bulgaria, reportedly, is in the process of removing its VAT accounts system.

- The *compulsory use of a third party* to guarantee VAT payments, either in general or for particular sectors, as set out by Ainsworth (2006), who labels this “D-VAT” (for “digital”).

While this may provide some assurance to the authorities, it is not clear that it solves the underlying problem. It seems, rather, to simply pass it to the guaranteeing

agency—which will presumably require to be compensated for the risk it bears, whether in the form of charges to government or to the taxpayers themselves.

These schemes, it should be noted, would address not only fraud related to export zero-rating—though this is certainly a key motivation—but also forms of purely domestic fraud arising from simple failure to remit tax due. The final administrative proposal is directed specifically to export-zero rating:

- Under the “*P-VAT*” of Poddar and Hutton (2001), goods would simply not be allowed to clear customs until the authorities had received confirmation (or an appropriate guarantee) that import VAT has been paid.
- While this to a large degree addresses the problem, it does so by reintroducing or strengthening restrictions on the movement of goods of precisely the kind that the EU seeks to remove.

Administrative solutions, thus, all have their weaknesses, whether in creating other opportunities for fraud or/and in increasing the compliance costs of honest traders. More radical proposals go to the heart of the problem by fixing the break in the VAT chain as goods pass between countries (or states within a federation): that is, they eliminate the zero-rating of exports. There are three main proposals of this kind:

- When the likely difficulties created for the VAT by the opening of borders within the EU first became apparent, the European Commission suggested a system of ‘exporter-rating’: intra-Union exports would bear VAT at the rate of the exporting country, with full credit then provided to the importer.²⁵ To ensure the same allocation of revenue as under zero-rating, a corresponding transfer would be made by the authorities of the exporting country to those of the importing country. Under the initial proposal, this adjustment would have been on the basis of individual transactions: a “clearing house” system. The alternative was later raised of allocating revenues on the basis of aggregate consumption data.

Apart from resistance to the expansion of the Union’s bureaucracy that the clearing house would seem to imply, the main weakness of this approach is in the poor incentives it creates for member states to monitor VAT payments (Lee, Pearson, and Smith, 1988). An importing country, for example, has little interest in verifying claims for credit on imports from other member states (or helping the exporting country to do so), since the cost of that credit (and hence of fraudulent refund claims) is simply passed to the exporting country. And if revenue is simply allocated in

²⁵ Confusingly, this is sometimes referred to in EU discussion as an “origin” form of VAT. It is not, however, an origin-based tax in the sense (more usual in the economics literature) set out in Box 1 above: that would require the importing country to give a credit at its own rate of VAT, not that of the exporting country. Exporter-rating is simply a means of implementing the destination principle, with commodities ultimately taxed at the rate of, and the revenue accruing to, the country in which they are purchased.

proportion to aggregate consumption, a member state would keep much less than €1 for each additional Euro of VAT that it collects.²⁶

- Under the “CVAT”—proposed by Varsano (1999) and further developed by McLure (2000)—exports would be formally zero-rated in the exporting country but then immediately subjected to a special ‘compensating’ VAT that would be fully creditable in the importing country. Thus the net revenue collected by this tax would in principle be zero (leaving aside complications created by exemptions and sales to final consumers).

This is an elegant scheme, fixing the VAT chain while at the same time preserving the allocation of revenues associated with zero-rating. Its main drawback is that exports continue to be treated differently from domestic sales. This runs counter to the objective of creating a unified single market, in the sense that a trader in Munich, for instance, would still need to treat a sale to Berlin (fully taxable) differently from one to Paris (zero-rated under the German VAT, but with the CVAT applicable).

- The “VIVAT”—developed by Keen and Smith (1996, 2000)—instead charges a single rate—common to all participating countries (or states)—on all B2B sales, whether domestic or across borders, while leaving the rate applied to final sales to the discretion of each. Under such a scheme, a trader in Munich would treat customers in Berlin and Paris in exactly the same way: so long as they are registered for VAT, she would charge the EU-wide B2B rate.

The VIVAT has been criticized on the grounds that it requires taxpayers to know whether or not their customer is registered for VAT. But this is already required, for example, for intra-Union exports of goods (zero-rating being available only for sales to registered taxpayers), and for reverse-charged supplies of services. Proposals for dealing with e-commerce also commonly envisage a distinction between B2B and other transactions.²⁷ A deeper concern again relates to the revenue allocation issue, since simple application of a common rate on B2B exports will increase revenue for net exporters of intermediate goods (and reduce revenue for net importers). Thus some form of clearing may again be needed.

There has been a lively debate on the relative merits of these schemes.²⁸ This need not be pursued here, since any of these more radical schemes would clearly offer a more robust defense against trade-related VAT fraud within the EU than do administrative solutions of the kind described above. But they are, nevertheless, radical, raising sensitive issues of tax

²⁶ A transactions-based clearing house of this kind has nevertheless operated, apparently with some success, between Israel and West Bank Gaza. But circumstances there, including not least the sheer volume of trade, are very different from those of the EU.

²⁷ See for example OECD (2001).

²⁸ See for instance the symposium in the December 2000 issue of *International Tax and Public Finance*.

sovereignty: Who, for example, would determine the rate of the CVAT, or of the intermediate rate under the VIVAT? Who would administer the CVAT? In combating fraud, however, they, in principle, create scope for all participating countries to gain. More fundamentally, they address the difficulties created by export zero-rating by pursuing further, not deviating from, the basic logic and inherent simplicity of the VAT.

VI. CONCLUSIONS—WITH LESSONS FOR THE UNITED STATES

There is no doubt that the VAT is susceptible to evasion and fraud, running all the way from the occasional concealed sale to sophisticated and large scale attacks by organized crime. Nor is there any doubt that, although many of the frauds to which the VAT is vulnerable also arise under, for example, a retail sales tax, the structure of the tax—notably the payment of refunds, particularly on exports—does create distinctive control problems. Certainly non-compliance with the VAT, particularly in the form of carousel fraud, is currently much on the minds of many policymakers, especially in the EU. And, making for dramatic stories, it is also much in the headlines.

But none of this is cause for panic. All taxes face problems of noncompliance. Some of these take a particularly striking (and, for those so inclined, tempting) form in the case of the VAT, with fraud potentially leading not merely to an inappropriate reduction in tax liability towards zero but outright cash payment to the fraudster. It is important to remember, however, that from the wider perspective of the government's finances \$1 paid in a fraudulent VAT refund is no more costly than \$1 of under-declared income tax. The question is not whether the VAT is vulnerable to fraud and evasion—it obviously is—but whether it is more or less vulnerable than other taxes. For, as discussed above, the fractional nature and self-recovering features of the VAT mean that in revenue terms it should in some respects be relatively robust to non-compliance: it could be that these effects, though not grabbing many headlines, more than offset more spectacular abuses. Some of the most careful estimates of the VAT gap, for the United Kingdom, put it currently at around 13.5 percent or, in a bad year, nearly 17 percent. This compares with a personal income tax gap likely to be somewhat less than the 14 percent found in the United States. Noncompliance with the VAT may in this sense now be more extensive in the United Kingdom than is noncompliance with the personal income tax—but the two seem likely to be of broadly the same order of magnitude.

The implications of these figures for the United States are of particular interest. For there is no reason to suppose that a federal VAT—if designed in line with current best practice (preferably with a single rate, reasonably high threshold and limited exemptions) and set in an appropriate administrative framework—would be any more vulnerable to noncompliance than that of the United Kingdom. Indeed in several respects a federal VAT in the United States should be less vulnerable: the far lower ratio of exports to GDP (7 percent compared to 17 percent in the United Kingdom, in 2005) implies a relatively low structural need for refunds; the particular difficulties that arise from extensive domestic zero-rating can be avoided by good design of the VAT; and, of course, the distinctive problems created by the absence of border controls and continuance of export zero-rating between member states—which are the focus of current concerns in the EU and discussed at some length

above—would simply not arise with a federal VAT in the United States so long as fiscal controls are maintained on trade with the rest of the world.

There is another reason not to panic. In a fundamental sense, there is no mystery as to how one might solve the current difficulties of carousel and related frauds in the EU. The deep solution is well-known: to fix the VAT chain by ending the zero-rating of trade between member states. And schemes to do this—notably the CVAT and VIVAT described above—are now conceptually reasonably well worked-out. Adoption of such schemes in the EU would be a major development, and their time may not yet have come: with 25 member states, the unanimity that the Union requires for all common tax measures creates huge inertia in dealing with coordination problems of this kind. Again, however, the implication for the United States deserves emphasis, and is by no means discouraging. Any VAT that endows the states with some discretion over rates or base would face a problem exactly analogous to that which perturbs policymakers in the EU: how to deal with inter-state trade, given the potential for evasion and fraud implied by the absence of internal border controls. When Europe adopted the VAT as a common form of general sales taxation it did not recognize that this problem would arise as regional integration deepened, and nor did it possess the intellectual apparatus to deal with it. If it were determined to decentralize some VAT design powers to the states, however, the United States, would need to confront the problem at the outset, and, happily, now has a range of potential solutions to consider.²⁹

None of this is intended to suggest that VAT fraud is an insignificant issue. Indeed the spectacular and highly organized nature of many carousel and related frauds is cause for particular concern, beyond the associated revenue loss itself, in the risk that it carries for the wider respect in which the tax system is held. But the serious current difficulties should not distract attention from the many aspects of the VAT compliance that appear to work well, and that may be quantitatively no less important. The current difficulties are, moreover, ones that, we believe, can with sufficient will be brought under control in the EU, and which the United States could, to a large degree, avoid if it were to move to a VAT, whether purely federal or with a state-level component too.

²⁹ See, for example, Keen (2001). Bird, Mintz, and Wilson (2006) argue that a federal VAT could readily co-exist with state level RSTs.

References

- Agha, Ali, and Jonathan Haughton, 1996, "Designing VAT Systems: Some Efficiency Considerations," *The Review of Economics and Statistics* 78, No. 2, pp. 303–08.
- Ainsworth, Richard T., 2006, "Carousel Fraud in the EU: A Digital VAT Solution," *Tax Notes International*, pp. 443–48.
- Aizenman, Joshua, and Yothin Jinjirak, 2005, "The Collection Efficiency of the VAT: Theory and International Evidence," NBER Working Paper No. 11539 (Cambridge, Massachusetts: National Bureau of Economic Research).
- Bird, Richard M., 1993, "Review of 'Principles and Practice of Value Added Taxation: Lessons for Developing Countries,'" *Canadian Tax Journal*, 41 No. 6, pp. 1222–25.
- _____, Jack M. Mintz, and Thomas A. Wilson, "Coordinating Federal and Provincial Sales Taxes: Lessons from the Canadian Experience," *National Tax Journal*, forthcoming.
- Boadway, Robin, Maurice Marchand, and Pierre Pestieau, 1994, "Towards a Theory of the Direct-Indirect Tax Mix," *Journal of Public Economics* 55, pp. 71–88.
- _____, and Pierre Pestieau, 2003, "Indirect Taxation and Redistribution: The Scope of the Atkinson-Stiglitz Theorem," In *Economics for An Imperfect World: Essays in Honor of Joseph E. Stiglitz*, ed. by Richard Arnott, Bruce Greenwald, Ravi Kanbur and Barry Nalebuff (Cambridge, Massachusetts: MIT Press) pp. 387–403.
- Brondolo, John, and Carlos Silvani, 1993, "An Analysis of VAT Compliance," mimeo (Washington: International Monetary Fund).
- Browning, Martin, and Costas Meghir, 1991, "The Effects of Male and Female Labor Supply on Commodity Demands," *Econometrica* 59, pp. 925–951.
- Cnossen, Sijbren, 1990, "Taxing Value Added: The OECD Experience," *International VAT Monitor* 5 (May), pp. 2–16.
- Commission of the European Communities, 2004, *Report from the Commission to the Council and the European Parliament on the Use of Administrative Cooperation Arrangements in the Fight Against VAT Fraud*, 260 final. COM
- Dharmapala, Dhammika, and Joel Slemrod, 2006, "Optimal Tax Remittance by Heterogeneous Firms," mimeo (Ann Arbor, Michigan: University of Michigan).
- Ebrill, Liam, Michael Keen, Jean-Paul Bodin, and Victoria Summers, 2001, *The Modern VAT* (Washington: International Monetary Fund).

- Fedeli, Silvia, and Francesco Forte, 1999, "Joint Income Tax and VAT-Chain Evasion," *European Journal of Political Economy* 15, pp. 391–415.
- Fox, William, and Matthew N. Murray, 2004, "Sales Taxation in a Global Economy," In *Taxing the Hard-to-Tax*, ed. by James Alm, Jorge Martinez-Vasquez, and Sally Wallace (Amsterdam: Elsevier), pp. 221–44.
- Harrison, Graham, and Russell Krelove, 2005, "VAT Refunds: A Review of Country Experience," IMF Working Paper 05/218 (Washington: International Monetary Fund).
- Hemming, Richard and John Kay, 1981, "The United Kingdom," In *The Value Added Tax: Lessons from Europe*, ed. by H. J. Aaron (Washington: Brookings Institution).
- Internal Revenue Service, 2002, "Compliance Estimates for Earned Income Tax Credit Claimed on 1999 Returns" (Washington: Internal Revenue Service).
- International Tax Dialogue, 2004, "The Value Added Tax: Experiences and Issues," available at www.itdweb.org
- Kay, John, 1989, "The Economic Functions of the Tax System," In *The Economic Borders of the State*, ed. by Dieter Helm (Oxford: Oxford University Press), pp. 218–236.
- Keen, Michael, 2001, "States' Rights and the Value Added Tax: How a VIVAT Would Work in the U.S." *Proceedings of the National Tax Association*, pp. 195–200.
- _____, 2006, "VAT, Tariffs and Withholding: Tax Design and the Informal Sector," mimeo (Washington: International Monetary Fund).
- Keen, Michael and Ben Lockwood, 2006a, "The Value Added Tax: Its Causes and Consequences," mimeo (London: University of Warwick).
- _____, 2006b, "Is the VAT a Money Machine?" mimeo (London: University of Warwick).
- Keen, Michael, and Jack Mintz, 2004, "The Optimal Threshold for a Value-Added Tax," *Journal of Public Economics* 88, pp. 559–76.
- Keen, Michael, and Stephen Smith, 1996, "The Future of Value-Added Tax in the European Union," *Economic Policy* 23, pp. 373–411 and 419–20.
- _____, 2000, "Viva VIVAT!," *International Tax and Public Finance* 7, pp. 741–51.
- Lee, Catherine, Mark Pearson, and Stephen Smith, 1988, *Fiscal Harmonisation: An Analysis of the European Commission's Recommendations*, Report Series No. 28 (London: Institute for Fiscal Studies).

- Matthews, Kent, 2003, "VAT Evasion and VAT Avoidance: Is There an European Laffer Curve for VAT?" *International Review of Applied Economics* 17, No. 1, pp. 105–114.
- _____, and Jean Lloyd-Williams, 2000, "Have VAT Rates Reached Their Limit? An Empirical Note," *Applied Economics Letters* 7, pp. 111–115.
- McLure, Charles E. Jr., 2000, "Implementing Sub-National Value Added Taxes on Internal Trade: The Compensating VAT (CVAT)," *International Tax and Public Finance* 7 pp. 723–40.
- Organisation for Economic Co-operation and Development, 2001, *Taxation and Electronic Commerce* (Paris: Organisation for Economic Co-operation and Development).
- Pedone, A., 1981, "Italy," In *The Value Added Tax: Lessons from Europe*, ed. by Henry J. Aaron (Washington: The Brookings Institution).
- Poddar, Satya, and Eric Hutton, 2001, "Zero-Rating of Inter-State Sales Under a Sub-National VAT: A New Approach," mimeo (Toronto: Ernst and Young).
- Ruffles, David, Geoff Tily, David Caplan, and Sandra Tudor, 2003, "VAT Missing Trader Intracommunity Fraud; The Effect on Balance of Payments Statistics and U.K. National Accounts," *Economic Trends* No. 597, pp. 58–70.
- Sinn, Hans-Werner, Andrea Gebauer, and Rüdiger Parsche, 2004, "The Ifo Institute's Model for Reducing VAT Fraud: Payment First, Refund Later," *CESifo Forum* 2, pp. 30–4.
- Slemrod, Joel, 2006, "Cheating Ourselves: The Economics of Tax Evasion," *Journal of Economic Perspectives*, forthcoming.
- Varsano, Ricardo, 2000, "Sub-National Taxation and Treatment of Interstate Trade in Brazil: Problems and a Proposed Solution," In *Decentralization and Accountability of the Public Sector*, ed. by Shahid Javed Burki, Guillermo Perry, and Others, Proceedings of the Annual Bank Conference on Development in Latin America and the Caribbean, (Washington: The World Bank) pp. 339–55.
- Zee, Howell, 2005, "Simple Analytics of Setting the Optimal VAT Threshold," *De Economist*, 153, pp. 461–71.
- _____, 2006, "VAT Treatment of Financial Services: A Primer on Conceptual Issues and Country Practices," *INTERTAX*, 34, pp. 458–74.

Official publications

Bank of England, *Inflation Report* (August, 2006) online at <http://www.bankofengland.co.uk/publications/inflationreport/ir06aug.pdf>

European Commission, *Report from the Commission to the Council and the European Parliament on the Use of Administrative Cooperation Arrangements in the Fight Against VAT Fraud*. COM (2004) 260 final.

U.K., Government. HM Customs and Excise. *Measuring Indirect Tax Losses* (November, 2002) online at <http://www.hm-treasury.gov.uk/media/389/E5/admeas02-297kb.pdf>

_____, *Measuring and Tackling Indirect Tax Losses. An update on the Government's Strategic Approach* (December, 2003) online at http://customs.hmrc.gov.uk/channelsPortalWebApp/downloadFile?contentID=HMCE_PROD_011582

_____, HM Revenue and Customs. *Measuring Indirect Tax Losses—2005* (December, 2005) online at <http://www.hmrc.gov.uk/pbr2005/mitl2005.pdf>

_____, National Audit Office. *HM Customs and Excise. Tackling VAT Fraud*. Report by the Comptroller and Auditor General, HC 357 Session 2003-2004 (3 March, 2004).