

# Developing Methodologies for Measuring Direct Tax Losses

---

October 2007

# Contents

---

1. Introduction	3
2. Methodologies	5
3. Estimates of Direct Tax Losses	12
<i>Individuals subject to Self-Assessment</i>	12
<i>Small and Medium-sized Employers</i>	13
<i>Small and Medium Sized Companies</i>	14
<i>Annex A: Detail for Individuals Estimates</i>	16
<i>Annex B: Detail for Small and Medium Employer Estimates</i>	21
<i>Annex C: Detail for Small and Medium Company Estimates</i>	24
<i>Glossary</i>	27

# 1 Introduction

---

1.1. This document discusses the methodological issues associated with estimating direct tax losses, sets out HMRC's approach to further development of direct tax loss measures and describes the more robust results available.

1.2. Estimating the scale of, and trends in, direct revenue losses is inherently difficult and is not widely attempted by governments around the world. No tax authority routinely publishes comprehensive tax loss estimates, and HMRC is at the forefront of work to develop estimates for a range of taxes and behaviours.

1.3. In general, two approaches can be employed to assess the scale of revenue lost from the tax system:

- **top-down method.** Estimates of revenue loss are produced indirectly by constructing an estimate of the total tax liability from sources independent of data collected for tax purposes, which is then compared with actual receipts;
- **bottom-up method.** Estimates of revenue loss are produced directly from departmental sources, such as surveys, administrative and operational data.

1.4. In principle, the top-down method is better as it is independent of both HMRC and taxpayers. In addition, it is more comprehensive because the estimates should cover all sources of losses. However, the top-down method requires sufficient data to be able to construct an independent estimate of tax liabilities. This is not always possible.

1.5. In contrast the bottom-up method is less comprehensive because by their nature many of the losses arise from activities that are deliberately concealed, and so are not covered by HMRC data sources.

1.6. The main sources of information on income are data collected for tax purposes so it is not possible to produce an independent measure of theoretical liability for direct taxes. Bottom-up measures are therefore being developed.

1.7. HMRC welcomes feedback on this paper from business, academic experts and other interested parties, with a view to discussing and, where possible, improving these methodologies further. To comment, or to request further copies of this paper, please use the following contact details.

Developing Methodologies For Measuring Direct Tax  
Losses.

KAI

Floor 3

HM Revenue and Customs

Somerset House

Strand

London

WC2R 1LB

020 7438 4355

## 2. Methodologies

---

2.1. For indirect taxes, where HMRC have independent sources of data on spending and the structure of the taxes is relatively simple, it is possible to employ a top-down method.

2.2. For direct taxes there are no independent sources of data describing the taxable income for the various taxpayer populations. Moreover, the structures of the taxes are more complex and a range of tax rates may apply to different parts of taxpayers' incomes. As a result the only approach open to HMRC is the bottom-up method.

2.3. In this method an estimate of revenue loss is produced by using information acquired from HMRC compliance activities. Most compliance work is risk based, and it can be difficult to use the information gained from such enquiries to assess losses from other taxpayers. HMRC also undertakes some random enquiries, the results of which can be used to extrapolate figures for the rest of the population.

2.4. This section discusses the issues surrounding the methodologies currently employed in some areas of direct taxes, and sets out other possible approaches.

### **Random Enquiries**

#### **Principle**

2.5. Random enquiries are the technique most commonly used by fiscal authorities, particularly in the USA, to measure direct tax losses arising from inaccurate reporting of income. Random enquiry programmes are tightly controlled statistical exercises, where samples of taxpayers are selected at random and their returns subjected to full enquiries by HMRC officers.

2.6. HMRC has separate random enquiry programmes for Income Tax Self Assessment (ITSA), Corporation Tax Self Assessment (CTSA) and Employers. Full details of the populations covered by these programmes are given in the annexes.

2.7. HMRC's random enquiry programmes provide results on the proportion of taxpayers under-reporting their tax liabilities, and the corresponding amount of the additional tax due. The results can be used to produce a figure for tax losses for the whole population because the enquiries are randomly selected and form a representative sample.

## Issues

2.8. A number of issues arise from the use of random enquiry techniques to measure tax losses. These include statistical uncertainty and the timing of the availability of results.

2.9. The estimates presented in this document represent HMRC's best current assessment of the levels of tax loss in certain areas of direct tax. However the true values may differ from these due to statistical variation and systematic uncertainty. This is mainly caused by two factors:

- sampling variation in sample data. The whole population is not subject to enquiry, so even though the sample is designed to be representative its characteristics may differ from the population purely by chance;
- systematic biases where the sample results consistently tend to underreport the true values for the population, or where the sample does not include the full population, for example those participating in the informal economy.

2.10. The random enquiry programmes will not identify all incorrect returns or the full scale of losses, especially where independent information from third parties is not available to verify the data supplied by the taxpayer. This means that tax loss estimates produced through random enquiries underestimate the full extent of the losses. Further, the proportion of losses that are not identified may also vary over time due to changes in the enquiry regime or the nature of non-compliance.

2.11. The Internal Revenue Service (IRS) in the USA has tackled this problem by using a range of 'multipliers' to make adjustments for non-detection<sup>1</sup>. These multipliers are generated through supplementary studies on particular tax return entries, together with econometric analysis of non-detection rates across IRS examiners. This work shows that undetected losses in some customer groups (particularly the self-employed) form a significant proportion of the total tax losses in the USA.

2.12. The IRS is able to undertake this analysis of non-detection because their random enquiry samples cover upward of 50,000 cases – much higher than is feasible in the UK. HMRC is continuing to explore ways in which the type of modelling carried out by the IRS could be deployed within the constraint of smaller sample sizes.

2.13. A further issue is the delay between the relevant tax year and the availability of enquiry results. This is an inevitable consequence of

---

<sup>1</sup> James Andreoni, Brian Erard and Jonathan Feinstein (1998) 'Tax Compliance', Journal of Economic Literature, Vol. 36, No. 2. (Jun., 1998), pp. 818-860

an enquiry-based approach. For example, individuals generally have until 31 January following the year of assessment to which the return relates to submit their tax return. HMRC then has a further year in which to open an enquiry. Particularly complex cases, which are often those involving large tax losses, can take a number of years to resolve. So it can be several years from the issue of a return before results are available for most of the sample.

2.14. Moreover, it remains necessary to make a forecast of the aggregate results of cases yet to be settled. The forecasts for such cases are based on the actual results of long running cases from previous years. Once actual results are known the losses figures are revised, which means the figures for losses reported in this document are subject to revision.

## **Areas for Further Research**

2.15. HMRC is also developing random enquiry programmes for other taxes, specifically Inheritance Tax (IHT) and Stamp Duty Land Tax (SDLT). However these programmes are at an early stage of development and it is not currently possible to produce robust estimates of tax losses from them.

2.16. It is not possible to use random enquiries to measure all types of loss or for all types of customer. The most significant types of losses that cannot be covered by random enquiries are:

- non-payment;
- large business;
- the use of avoidance schemes/devices to reduce liability;
- the informal economy; and
- individuals who are not issued a return.

2.17. Losses from non-payment can be measured from HMRC's accounting systems, but the other types of loss are much more difficult to measure. Some of the methodological issues for these areas are set out below.

### **Large Business**

2.18. A full enquiry can involve a complete examination of all books and records. For the largest businesses this is not a practical approach because of the huge resource requirement it would place on both HMRC and the business. However HMRC carries out regular risk assessments for all large businesses to identify the key areas where

tax treatment is in doubt. These risk assessments are recorded on a new management information system. HMRC is planning to use this information to inform understanding of the size of losses for large businesses.

2.19. There are a number of issues with estimating tax losses for the whole population in this way:

- the figure recorded needs to be regarded with caution as initially it is made subjectively by individual tax specialists without full knowledge of the business and its circumstances;
- the risk assessments look at the potential tax under consideration. This may well change over time as a result of engagement with the company (and possibly settled through litigation); and
- some issues could be missed in risk assessment.

2.20. Also as it takes forward the recommendations from the Varney Review of Links with Business, HMRC is increasingly working with large businesses in real time to address potential risks at an early stage.

2.21. These issues are being addressed in the design of the new processes and databases introduced in 2006 to record risks. As the quality of the data improves over time, HMRC should be in a better position to determine the trend in tax losses for this population.

2.22. One suggested proxy for Corporation Tax losses for large business is to use the information provided by the companies themselves in publicly available accounts in order to calculate an “effective tax rate” (the ratio of tax actually paid to commercial profits), which could be compared with the statutory tax rate in order to estimate losses.

2.23. Other countries have used variations in effective tax rates as an indicator of tax performance. These methods are more difficult to centrally apply in a UK context because there is no group basis of company taxation (unlike many other leading OECD countries) and our accounting standards make it more difficult to identify the UK element of quoted companies’ commercial profits.

## **Avoidance**

2.24. Tax losses arising from the use of avoidance schemes/devices cannot be measured solely through the results of enquiries.

2.25. The avoidance disclosure regime was first introduced in 2004 and initially covered Corporation Tax (CT) financial, Income Tax (IT) on

employment income<sup>2</sup> and Capital Gains Tax (CGT). Since then the regime has been extended to the rest of CT and the rest of IT in August 2006, with Stamp Duty Land Tax (SDLT) included in August 2005 and National Insurance contributions (NICs) in August 2007.

2.26. As part of the regime, promoters of schemes have to disclose to HMRC the details of the schemes developed and users should notify HMRC of a scheme in their tax return. This is starting to provide a greatly enhanced view of the use of marketed avoidance schemes and the potential associated losses. A full publication schedule of quarterly information at six monthly publication dates has been in place for nearly 18 months.

## **Informal Economy**

2.27. The informal economy relates to businesses and individuals who remain outside the tax system by not declaring their economic activity or tax affairs to HMRC. They are not accurately recorded by any government agency or survey. Since these individuals do not appear on HMRC's systems it is not possible to measure the extent of the problem through an enquiry. Therefore a different kind of approach is required.

2.28. Measurement of the size of the informal economy receives a lot of attention amongst academics and there have been a number of attempts to produce models to estimate its size. HMRC continually reviews this work for developments that will help in the calculation of tax losses.

2.29. HMRC has done work to compare numbers of people in work from available surveys with numbers shown on HMRC databases. These have not proved to be useful comparisons and have not shown any patterns thought to provide useable indicators of levels of non-compliance.

2.30. The problem is that there is too much statistical noise in survey data to allow proper comparisons; and there are questions over the extent to which genuine participants in the informal economy would take part in surveys. This reluctance to admit to participation in the informal economy has been confirmed in a recent HMRC study to explore the feasibility of using dedicated survey approaches in this area.

---

<sup>2</sup> Products/schemes that aim to avoid tax on employment income.

2.31. Econometric methods include examining measures like the circulation of large denomination bank notes, which could indicate the scale of activity in the informal economy. But these suffer from several problems, including:

- it is not possible to distinguish the proportion of illegitimate cash or cheque payments due to informal work and the proportion due to crime, or how this changes over time;
- while these studies may help identify changes in the informal economy it is hard to use them to measure its level; and
- cash and cheque payments are influenced by a number of factors other than the informal economy.

2.32. A recent OECD report<sup>3</sup> concluded that *'The problems with these models are evidenced by the sensitivity of the results to the benchmark year assumptions, the wide range of estimates given under the same circumstances, and the implausible results that are obtained in some specific circumstances'*.

2.33. In the absence of other types of approach HMRC is concentrating on enhancing its ability to detect people operating in the informal economy - using an increased capacity to compare HMRC data to a wider set of third party sources. Regular comparison exercises against a common set of information sources are expected to provide an improved picture of the informal economy and trends within it.

### **Individuals who are not issued a return**

2.34. It is not necessary to issue most taxpayers with a return where all tax liabilities are withheld at source, for example employment income, where tax is deducted under PAYE, or basic rate tax withheld from bank interest.

2.35. It is not possible for HMRC to conduct random enquiries into the tax affairs of individuals who are not issued a return because the legal position requires a return to be issued for an enquiry to take place. In general, HMRC believes that the tax at risk for individuals who do not receive a tax return is low. This is because most are employees who pay the bulk of their tax through PAYE and as such, the opportunities for non-compliance are limited<sup>4</sup>. Therefore the burden that this would impose on largely compliant taxpayers would be disproportionate.

---

<sup>3</sup> OECD (2002) Measuring the Non-observed Economy: A Handbook, Paris: OECD. Available at [www.oecd.org](http://www.oecd.org)

<sup>4</sup> The IRS (2007: 12) notes [*'Reducing the Federal Tax Gap: A Report on Improving Voluntary Compliance'*. 2007. Internal Revenue Service & U.S. Department of the Treasury, downloaded from [http://www.irs.gov/pub/irs-news/tax\\_gap\\_report\\_final\\_080207\\_linked.pdf](http://www.irs.gov/pub/irs-news/tax_gap_report_final_080207_linked.pdf) on 2 October 2007] *'The amounts least likely to be misreported on tax returns are subject to third-party information reporting and withholding and are, therefore, the most visible (e.g. wages and salaries). The net misreporting percentage for wages and salaries is only 1.2 percent.'*

2.36. Nevertheless, there are risks in this population. Taxpayers may not inform HMRC about sources of income such as rental income or about capital gains, where there is no withholding or automatic third party information reporting. For instance, the growth of the buy-to-let business in the property market has increased the risk of undeclared rental income and capital gains.

2.37. To tackle these risks, and to obtain a better view of the extent of losses, HMRC is developing improved risk assessment techniques for identifying undeclared sources of income. This approach is very similar to that being developed for the informal economy.

## **Conclusion**

2.38. HMRC is attempting to develop a range of measurement techniques that cover all types of loss across customer groups and all taxes. However, the preceding discussion shows that the measurement of direct tax losses is extremely difficult and this is reflected by the fact that no tax authority routinely publishes comprehensive direct tax loss estimates. Progress is being made but there are a number of key areas where HMRC does not yet possess robust methodologies.

2.39. HMRC does have confidence in the figures that come from random enquiry programmes and analysis of non-payment. The rest of this report describes the results of random enquiries for individuals, and small and medium employers and companies together with non-payment figures for these groups. Details of the methodologies employed are shown in:

- Annex A for Individuals;
- Annex B for Small and Medium Employers;
- Annex C for Small and Medium Companies.

2.40. Quantifying tax losses is important in helping to improve our understanding of such losses and will ultimately improve revenue collection. Therefore HMRC is continuing to develop methodologies as access to more reliable and wider sources of data and sophisticated modelling techniques become available. It is the intention of HMRC to publish further estimates of tax losses when robust methodologies have been put in place.

## 3. Estimates of Direct Tax Losses

3.1. This section covers losses from non-payment and incorrect returns for individuals (including the self-employed, small partnerships and trusts), small and medium-sized employers, and small and medium-sized companies.

3.2. The results presented here come from random enquiries and non-payment data.

### **Individuals subject to Self-Assessment**

3.3. Estimates of losses in this section are taken from the Income Tax Self Assessment (ITSA) Random Enquiry Programme and non-payment data, covering individuals who receive a return. In this context 'individuals' should be taken to mean individuals who are self-employed, and partnerships, as well as those who are employees or pensioners, or may only have investment income. Trusts also receive income tax returns and are included in the figures. The taxes covered are:

- Income tax;
- Capital gains tax; and
- Class 4 National Insurance Contributions (NICs).

3.4. Overall the tax and NICs collected through ITSA in 2001-02 represented about 7 per cent of total direct tax receipts for that year. Although individuals who do not receive a return are not included in this group, many will be employees taxed under Pay as You Earn (PAYE). Incorrect operation of PAYE schemes and consequent losses of tax due from employment earnings are dealt with in the 'Employer' section of this chapter.

### **Results**

3.5. The most recent results available are for the 2001-02 tax year.

**Table 3.1: Losses identified for ITSA returns (£ billion)<sup>2</sup>**

	1999-2000	2000-01	2001-02
Losses due to incorrect returns	2.4	2.7	2.8
Compliance Yield <sup>1</sup>	0.4	0.4	0.4
Non-payment	0.1	0.1	0.1
<b>Total tax losses identified</b>	2.1	2.3	2.5
<b>Total tax liabilities</b>	19.6	19.8	19.2
<b>Proportion of liabilities (per cent)</b>	11%	12%	13%

<sup>1</sup> Enquiries settled during each year.

<sup>2</sup> Figures rounded to the nearest £0.1bn. As a result components may not appear to sum.

3.6. Table 3.1 shows the estimated tax losses identified for 1999-00 to 2001-02. These figures do not adjust for non-detection. Over this

period the estimated losses increased slightly from 11 to 13 per cent of liabilities, though this change is within the confidence intervals.

## **Small and Medium-sized Employers**

3.7. Estimates of losses in this section come from the Employer Random Enquiry Programme and non-payment data. This includes income tax and National Insurance contributions (NICs).

3.8. In this context 'small and medium-sized employers' refers to those with up to 500 employees, except where the employer is part of a complex group. Receipts of tax and NICs collected through Pay As You Earn (PAYE) schemes from this customer group encompassed about 27 per cent of total direct tax receipts for 2003-04.

### **Results**

3.9. The most recent results available are for the 2003-04 tax year.

3.10. Table 3.2 shows identified losses of income tax and NICs for small and medium-sized employers were an estimated 1.2 per cent for 2003-04. This figure does not adjust for non-detection. The confidence interval for this estimate is 0.8 per cent to 1.6 per cent.

**Table 3.2: Losses identified for Small and Medium Sized Employers (£ billion)<sup>2</sup>**

	<b>2003-04</b>
Losses due to incorrect returns	0.9
Compliance Yield <sup>1</sup>	0.2
Non-payment	0.2
<b>Total tax losses indentified</b>	<b>0.8</b>
<b>Total tax liabilities</b>	<b>66.2</b>
<b>Proportion of liabilities (per cent)</b>	<b>1.2%</b>

<sup>1</sup> By period of settlement of enquiry.

<sup>2</sup> Figures rounded to the nearest £0.1bn. As a result components may not appear to sum.

## **Small and Medium-Sized Companies**

3.11. All companies chargeable to corporation tax are required to make returns of their profits under Corporation Tax Self Assessment (CTSA).

3.12. Estimates of losses from incorrect returns in this section come from the CTSA Random Enquiry Programme. In this context, 'small and medium-sized companies' means those companies not dealt with by HMRC's Large Business Service, which handles the largest and most complex companies. Some other companies, particularly other large ones identified by a combination of factors, such as trade, company type, turnover, international connections and capital base are

also not included. Receipts of tax from the companies covered in this section made up about 4 per cent of all direct tax receipts in 2001-02.

## Results

3.13. At present the most recent results available are for accounting periods ending in the year to 31 December 2003.

**Table 3.3: Tax losses identified for small and medium-sized companies (£ billion)<sup>2</sup>**

	2001	2002	2003
Losses due to incorrect returns	2.2	2.5	2.0
Compliance Yield <sup>1</sup>	0.4	0.4	0.4
Non-payment	0.0	0.0	0.1
<b>Total tax losses identified</b>	1.8	2.2	1.7
<b>Total tax liabilities</b>	12.6	13.3	13.6
<b>Proportion of liabilities (per cent)</b>	14%	17%	12%

<sup>1</sup> By period of settlement of enquiry. Financial year figures.

<sup>2</sup> Figures rounded to the nearest £0.1bn. As a result components may not appear to sum.

3.14. Table 3.3 shows estimated tax losses for calendar years 2001 to 2003. These figures do not adjust for non-detection. The proportion of losses as a percentage of the total tax liabilities has varied between 12 and 17 per cent. But allowing for confidence intervals the trend in these figures is broadly stable.

# Annex A: Detail for Individuals Estimates

---

## Methodology

### General Methodology

A.1. The estimate of direct tax losses resulting from non-payment and submission of incorrect returns in this section is limited to individuals who receive a tax return. In this context, 'individuals' means individuals who are self-employed, and partnerships (with up to 4 partners), as well as those who are employees or may only have investment income. So the taxes directly included are:

- Income tax;
- Capital gains tax; and
- Class 4 NICs.

A.2. Many individuals within Income Tax Self Assessment (ITSA), as well as many who do not receive a return, will be employees taxed under Pay as You Earn (PAYE). The tax payable through ITSA, as shown in Table 3.1, is therefore only a small proportion of the tax liabilities of this customer group.

A.3. This method will fail to identify the tax loss in the informal economy, and will not fully include losses from avoidance.

A.4. A bottom-up estimate is used, combining the results of the random enquiry programme with the amount of non-payment for this customer group. As some of the losses are recovered through HMRC compliance activity, this is subtracted to give the net losses. The loss estimate is defined by:

$$\text{Losses} = \text{Non-compliance} + \text{Non-payment} - \text{Yield from compliance activity}$$

A.5. The random enquiry programmes will not identify all incorrect returns or the full scale of losses, especially where independent information from third parties is not available to verify the data supplied by the taxpayer. This means that tax loss estimates produced through random enquiries underestimate the full extent of the losses.

A.6. The proportion of these losses as a percentage of all tax liabilities for this group is given by:

$$\text{Proportion of liabilities} = \frac{\text{Losses}}{\text{Liabilities}} \times 100$$

A.7. The total tax liability is total declared tax arising for the tax year (mostly income tax, but including capital gains tax and Class 4 NICs) plus the estimated losses from incorrect returns.

### ***The 1999-00 to 2001-02 Random Enquiry programmes***

A.8. The random enquiry program allows HMRC to estimate the losses arising from the submission of incorrect returns.

A.9. The random sample is selected from ITSA taxpayers issued with a notice to file a return.

A.10. The sample is drawn by a systematic process that selects every “n<sup>th</sup>” notice. The sampling interval, n, is determined by dividing the total number of returns issued by the required sample size (rounded down to the nearest whole number). When the return includes a partnership income schedule, the next return on the list is selected. This is because the returns of individuals who are partners will automatically be included in any enquiry resulting from the selection of a partnership return.

A.11. Sample sizes are given in Table A.1. Each return selected is subject to a full enquiry, which involves a complete examination of books and records. In exceptional circumstances, an enquiry is not carried out, but the caseworker must ensure that it meets internal criteria listed as ‘acceptable reasons’. The 95 per cent confidence intervals shown in Section 3 have been estimated using standard sampling theory.

**Table A.1: Sample sizes for the ITSA random enquiry programme**

<b>Tax Return Year</b>	<b>Sample size</b>
1999/00	6,700
2000/01	6,435
2001/02	6,434

A.12. Yield amounts are positive tax yield in favour of HMRC (excluding adjustments in the taxpayer’s favour) plus interest and penalties. Annual tax losses are calculated as yield (as defined above) less interest and penalties divided by the number of years covered by settlement.

### **Validation**

A.13. As part of each year’s programme, HMRC conducts a validation exercise for a sample of cases, focusing on those cases with the largest tax losses, which have the greatest influence on the results. The yield amounts are checked to confirm that the enquiry outcome has been recorded accurately.

## Timing

A.14. Estimates of tax losses from incorrect income tax returns rely on data from random enquiries. Individuals generally have until 31 January following the year of assessment to which the return relates to submit their return. HMRC then has a further year in which to open an enquiry. While many enquiries are dealt with quite quickly, others that are complex or reveal large tax losses take longer. It can be several years before results are available. For ITSA 2001-02 is the latest year for which results are available.

## On-going cases

A.15. In order to estimate losses for each year, it is necessary to make assumptions about the cases that were yet to be settled at the date the enquiry results are analysed. Forecasts of the results of such cases have been based on the results of similar settled cases. The forecasts are revised with the actual figures when available for future analyses. Therefore the figures for losses due to incorrect returns are subject to revision.

## Grossing methodology

A.16. To produce population estimates for total tax losses, the average loss estimates for the population are multiplied by the number of notices to deliver returns issued – excluding individuals who are partners, as they will be included in the population via the partnership return.

A.17. Adjustments are made for cases selected but dropped – that is, no enquiry is made because the return satisfies the criteria for not taking up an enquiry.

## ***Non-Payment***

A.18. The figures used to estimate levels of non-payment come from analysis of write-offs and remissions of tax and are collated for periods of approximately a year ending in October each year. They are then converted to financial year figures.

A.19. As separate figures of non-payment are not available for the taxpayers within the scope of this random enquiry programme, the amounts are apportioned by reference to the total tax liabilities of the ITSA customer group. These non-payment figures will tend to relate to tax liabilities for years before the random enquiry period. This approach has been taken because figures are not readily available by reference to the liability period.

## **Compliance Yield**

A.20. The random enquiries provide an estimate of the tax losses due to incorrect returns. However, HMRC carries out a programme of compliance activity to identify and correct incorrect returns. To calculate the net tax losses it is therefore necessary to subtract the yield from this activity. The figures for yield are taken from HMRC's systems for recording the outcomes of enquiries, and relate to cases settled during each year rather than enquiries into returns relating to a specific tax year.

## **Supplementary Results**

A.21. Table A.2 shows that for the tax years 1999-00 to 2001-02 the proportion of all ITSA returns that were estimated to be incorrect and leading to a loss of tax varied between 31 per cent and 33 per cent.

A.22. For between 7 per cent and 8 per cent of returns, the annualised level of under-declaration was over £1,000. However, in the remainder of cases where an under-declaration was found the figure was relatively small: between £1 and £500. Over 40 per cent of the under-declared liabilities relate to just two per cent of taxpayers.

**Table A.2: Incorrect ITSA returns with under-declared tax liability (Per Cent)**

	1999-2000	2000-01	2001-02
<b>Proportion</b>	31	32	33
<b>Of which, under-declared liability per annum</b>			
£1 to £500	17	17	18
£501 to £1,000	6	7	7
over £1,000	7	8	8

### Self employed

A.23. Any individual who has business income must complete the self-employment pages of the Income Tax Return. Partnerships also complete a partnership return showing the profits of their business, with the individual partners showing their share of the partnership income in the partnership pages of their own tax return.

A.24. Table A.3 shows that about half of the returns submitted by self-employed taxpayers under-stated the true tax liability. For about a quarter of returns the annualised understated liability was under £500, while for around 15 per cent it was over £1,000.

**Table A.3: Self employed taxpayers: Incorrect ITSA returns with under-declared tax liability (Per Cent)**

	1999-2000	2000-01	2001-02
<b>Proportion</b>	47	51	53
<b>Of which, under-declared liability per annum</b>			
£1 to £500	23	23	25
£501 to £1,000	11	14	12
over £1,000	14	14	16

Other individuals receiving a tax return

A.25. Individuals without business income who receive notices to deliver returns:

- company directors;
- other employees and pensioners taxed under PAYE, who are liable to tax at higher rates or who have relatively complicated tax affairs; and
- personal representatives or trustees.

The majority of income tax paid by these individuals is deducted at source under PAYE. The very low rates of loss in this group are shown in Table A.4.

**Table A.4: Other individuals receiving a return: Incorrect ITSA returns with under-declared tax liability (Per Cent)**

	1999-2000	2000-01	2001-02
<b>Proportion</b>	18	16	17
<b>Of which, under-declared liability per annum</b>			
£1 to £500	13	13	12
£501 to £1,000	2	2	2
over £1,000	2	2	3

A.26. Table A.4 shows that between 16 and 18 per cent of taxpayers in the “other individuals” category under-reported their true tax liability. For the majority of these (12–13 per cent of the total) the level of annualised under-reporting was under £500; while for 2 to 3 per cent the annual tax loss was over £1,000.

# Annex B: Detail for Small and Medium Employer Estimates

---

## Methodology

### General Methodology

B.1. The estimate of direct tax losses resulting from non-payment and submission of incorrect returns in this section relates to employers with 500 employees or fewer. We have termed these small and medium-sized employers. The employer may be a self-employed individual, partnership, or a company, and will be required to make returns under the PAYE Regulations and account for income tax and NICs.

B.2. Tax losses arising outside PAYE regulatory failure are unlikely to be picked up by random reviews. Different methodologies are being proposed for other losses, such as the shadow economy, and estimates will be published once sufficiently robust.

B.3. A bottom-up estimate is used, combining the results of the random enquiry programme with the amount of non-payment for this customer group. As some of the losses are recovered through HMRC compliance activity, this is subtracted to give the net losses. The loss estimate is defined by:

$$\text{Losses} = \text{Incorrect Returns} - \text{Compliance Yield} + \text{Non-Payment}$$

B.4. The random enquiry programmes will not identify all incorrect returns or the full scale of losses, especially where independent information from third parties is not available to verify the data supplied by the taxpayer. This means that tax loss estimates produced through random enquiries underestimate the full extent of the losses.

B.5. The proportion of these losses as a percentage of all tax liabilities for this group is given by:

$$\text{Proportion of liabilities} = \frac{\text{Losses}}{\text{Liabilities}} \times 100$$

B.6. The total tax liability is the tax declared for the tax year (income tax under PAYE, and employers' and employees' NICs) plus the estimated tax losses from incorrect returns.

## ***The 2003-04 Random Enquiry programme***

B.7. The Employer random enquiry programme allows us to estimate the losses arising from PAYE failures and other irregularities. Although the 2003-04 programme is the fourth random enquiry programme to be run on employers, changes in data available and methodology have meant that it is not possible to compare the results of previous years. For this reason, only the results of one year are presented.

B.8. Stratified random samples were created by selecting between 300 and 500 employers from each of 5 employer segments among employers with up to 500 employees. The segments are based on criteria such as number of employees. The total sample size for 2003/04 was 1,649 cases. The 95 per cent confidence intervals shown in Section 3 have been estimated using standard sampling theory.

### Validation

B.9. Each year HMRC conducts a validation exercise for a sample of cases. The yield amounts are checked to confirm that the enquiry outcome has been recorded accurately.

### Timing

B.10. Employer Compliance Reviews initially look at the records of the previous 12 months. While many enquiries are dealt with quite quickly, some may take longer where returns are complex or a high amount of under-declaration is found. Therefore it can be some years before final results are available. Here, 2003-04 is the latest year available.

### On-going cases

B.11. In order to estimate losses for each year, it is necessary to make assumptions about the cases that were yet to be settled at the date the review results are analysed. Forecasts of the results of such cases have been based on the results of similar settled cases. The forecasts are revised with the actual figures when available for future analyses. Therefore the figures for losses due to incorrect returns are subject to revision.

### Grossing methodology

B.12. Population estimates were derived from counts of the total number of PAYE schemes.

B.13. Adjustments are made for any cases selected but not taken up for review. A case has to fulfil certain criteria before it can be dropped.

## ***Non-Payment***

B.14. The figures used to estimate levels of non-payment come from analysis of remissions and write-offs of tax and are collated for periods of approximately a year ending in October each year. They are then converted to tax year figures.

B.15. Separate figures are not available for the small and medium-sized employers specifically within the scope of this random enquiry programme, so the non-payment estimates have been apportioned by reference to the total tax liabilities of employers. The non-payment figures will tend to relate to tax liabilities for years before the random enquiry period. This approach has been adopted because figures are not readily available by the liability period.

## ***Compliance Yield***

B.16. The random enquiries provide an estimate of the amount of the tax losses through PAYE and other irregularities by the employer. However HMRC carries out a programme of compliance activity to identify and correct such irregularities. To calculate the net tax losses it is therefore necessary to subtract the yield from this activity. The figures for yield from employer compliance reviews are taken from HMRC's systems for recording the outcomes of enquiries, and relate to cases settled during 2003-04 rather than reviews covering a particular liability period.

# Annex C: Detail for Small and Medium-Sized Company Estimates

---

## Methodology

### *General Methodology*

C.1 The estimate of direct tax losses resulting from non-payment and submission of incorrect returns in this section is limited to small and medium-sized companies. In this context, 'small and medium-sized' means those companies not dealt with by the Large Business Service (LBS) and also excludes the 10,000 or so largest and most complex companies dealt with outside the LBS.

C.2 A bottom-up estimate is used, combining the results of the random enquiry programme with the amount of non-payment for this customer group. As some of the losses are recovered through HMRC compliance activity, this is subtracted to give the net losses. The loss estimate is defined by:

$$\text{Losses} = \text{Incorrect Returns} - \text{Compliance Yield} + \text{Non-Payment}$$

C.3 The random enquiry programmes will not identify all incorrect returns or the full scale of losses, especially where independent information from third parties is not available to verify the data supplied by the taxpayer. This means that tax loss estimates produced through random enquiries underestimate the full extent of the losses.

C.4 The proportion of these losses as a percentage of all tax liabilities for this group is given by:

$$\text{Proportion of liabilities} = \frac{\text{Losses}}{\text{Liabilities}} \times 100$$

C.5 The theoretical tax liability is total tax liabilities for accounting periods ending in the calendar year (mostly corporation tax, but will include some income tax and tax charged under Section 419 ICTA 1988) plus the estimated losses from incorrect returns.

### *The 2001 to 2003 Random Enquiry programme*

C.6 The CT random enquiry programme allows us to estimate the losses arising from the submission of incorrect returns.

C.7 The random sample is selected from small and medium-sized companies issued a notice to deliver a return each month. The sampling process selects every  $n^{\text{th}}$  return.

C.8 The sample sizes are given in Table C.1. The programme covers companies with accounting periods ending in calendar years 2001 to 2003. The confidence intervals shown in Section 3 are calculated using a bootstrapping technique (which produces their non-symmetric shape (Figure 3.2)).

**Table C.1: Sample sizes for the CTSA random enquiry programme**

Accounting period ending in year	Sample size
2001	297
2002	315
2003	342

#### Validation

C.9 Each year HMRC conducts a validation exercise for a sample of cases, focusing on those cases with the largest tax losses, which have the greatest influence on the results. The additional liability amounts are checked to confirm that the enquiry outcome has been recorded accurately.

#### Timing

C.10 Companies have until a year after the end of their accounting period to submit a return and HMRC then has a further year in which to open an enquiry. Particularly in complex cases, the enquiry may not be settled until some years after that. For that reason, 2003 is the latest year for which analysis is available.

#### On-going cases

C.11 In order to estimate the overall losses for each year, it is necessary to make a forecast of the aggregate results of cases yet to be settled. The forecasts for such cases are based on the actual results of long running cases from previous years. Once actual results are known the losses figures are revised, which means the figures for losses reported in this document are subject to revision.

#### Grossing methodology

C.12 To produce population estimates for total tax losses, the average loss estimates from the enquiries are multiplied by the number of notices to deliver returns issued.

C.13 Adjustments are made for cases selected but where no enquiry was conducted – for example, because the company was missing or dormant.

## **Non-Payment**

C.14 The figures used to estimate levels of non-payment come from analysis of remissions and write-offs of tax and are collated for periods of approximately one year ending in October each year. They are then converted to calendar year figures.

C.15 Separate figures are not available for the small and medium-sized companies specifically within the scope of this random enquiry programme, so the non-payment estimates have been apportioned by reference to the total tax liabilities of this customer group. The non-payment figures will tend to relate to tax liabilities for years before the random enquiry period. This approach has been adopted because figures are not readily available by the liability period.

## **Compliance Yield**

C.16 The random enquiries provide an estimate of the tax losses due to incorrect returns. However, HMRC carries out a programme of compliance activity to identify and correct incorrect returns. To calculate the net tax losses it is therefore necessary to subtract the yield from this activity. The figures for yield are taken from HMRC's systems for recording the outcomes of enquiries, and relate to cases settled during each year rather than enquiries into returns relating to accounting periods ending in a specific year.

## **Supplementary results**

C.17 Table C.2 shows that for the years 2001 to 2003, the proportion of small and medium-sized companies submitting an incorrect return leading to a loss of tax was in the range 37 per cent to 43 per cent. This should be considered to be relatively constant in the context of the margin of error of these figures. The latest figure – for 2003 – relies more heavily on forecast results for ongoing cases and is therefore more sensitive to revisions as further enquiries are settled.

C.18 For accounting periods ending in 2001 and 2002, about a quarter of cases were settled with an annualised additional liability of over £1,000. The proportion appears a little lower for 2003 but could well increase as the results of more on-going cases are known.

**Table C.2: Small and medium-sized companies: Incorrect CTSA returns where additional tax liability established (per cent)**

<b>Accounting period ending in year</b>	<b>2001</b>	<b>2002</b>	<b>2003</b>
<b>Proportion</b>	43	43	37
<b>Of which, annualised additional liability</b>			
£1 to £1,000	16	18	18
over £1,000	28	25	19

# Glossary of technical terms

---

**Bootstrapping:** A statistical method that can be used to calculate confidence intervals using only the sample data as a distribution. It involves resampling with replacement, and is often used where the data set is small or particularly skewed.

**Compliance:** State of being in accordance with established legislation.

**Compliance activity:** An intervention by HMRC, such as a direct tax enquiry, employer compliance review or VAT assurance visit, design to ensure that the correct amount of tax is being accounted for and paid.

**Confidence Interval:** A range of values that has a specified probability of containing the true value of interest.

**Econometrics:** Combination of statistics and economic theories, principally used to analyse and test economic theories.

**Estimates:** Approximate result calculated from approximate or incomplete data.

**Evasion:** the deliberate omission, concealment or misrepresentation of information, or the false or deceptive presentation of circumstances, to fraudulently obtain a fiscal advantage, whether temporarily or permanently.

**Margins of Errors:** Measures of how reliable the estimated value is.

**PAYE Scheme:** Each employer operating PAYE registers a PAYE scheme with HMRC, which allows for the issue and monitoring of returns.

**Remissions:** Debts that HMRC is not pursuing.

**Sampling Weights:** In order to make the results from the sample more representative of the population, respondents' answers are "weighted" (multiplied by a coefficient) to rebalance the survey answers to the population profile.

**Self Assessment:** A system for reporting income and capital gains to, and claiming tax allowances from, HMRC. It involves completing a paper or online form.

**Settlement:** Closure of a direct tax enquiry, resulting in the agreement of any additional tax liability.

**Write offs:** Debts that are considered to be irrecoverable.

**Yield:** Money collected from compliance activity.

## ***Abbreviations***

**CGT:** Capital Gains Tax.

**CT:** Corporation Tax.

**CTSA:** Corporation Tax Self Assessment.

**EU:** European Union.

**HMRC:** Her Majesty's Revenue and Customs.

**IHT:** Inheritance Tax.

**IRS:** Internal Revenue Service.

**IT:** Income Tax.

**ITSA:** Income Tax Self Assessment.

**LBS:** Large Business Service.

**LFS:** Labour Force Survey.

**NAO:** National Audit Office.

**NICs:** National Insurance Contributions.

**ONS:** Office for National Statistics.

**PAYE:** Pay as You Earn.

**PBR:** Pre-Budget Report.

**SDLT:** Stamp Duty Land Tax.

**UK:** United Kingdom.