HM Revenue & Customs

Measuring tax gaps 2018 edition
Methodological annex
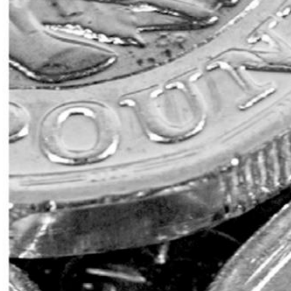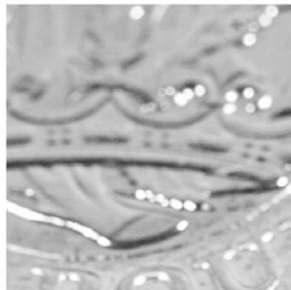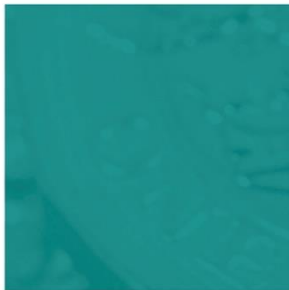
Contents

Chapter A: Introduction

A.1   This document provides further details of the data and methodology used to produce estimates of the tax gap published in 'Measuring tax gaps 2018 edition'. There are numerous approaches to measuring tax gaps. These can be categorised as:

- **data matching** – comparisons between related datasets

- **top-down methods** – these methods use external data sources to estimate total consumption of taxable products – used to calculate the total theoretical liabilities; the tax gap is the difference between the total theoretical liabilities and the tax received

- **random enquiries (bottom-up)** – these are full enquiries opened by HMRC officers into a randomly selected sample of taxpayers

- **management information (bottom-up)** – these methods use management information such as:

   o   risk registers (a list of identified tax risks, together with information such as estimated value, nature and status)

   o   data extracted from accounting systems

   o   other databases/systems used to manage HMRC's business

- **experimental** – where limited data is available, illustrative estimates are produced using assumptions made in collaboration with HMRC's operational experts.

A.2   The figure below shows the general approach used to calculate each tax gap component.

Approaches used to calculate the tax gap

Chapter B: Margins of error
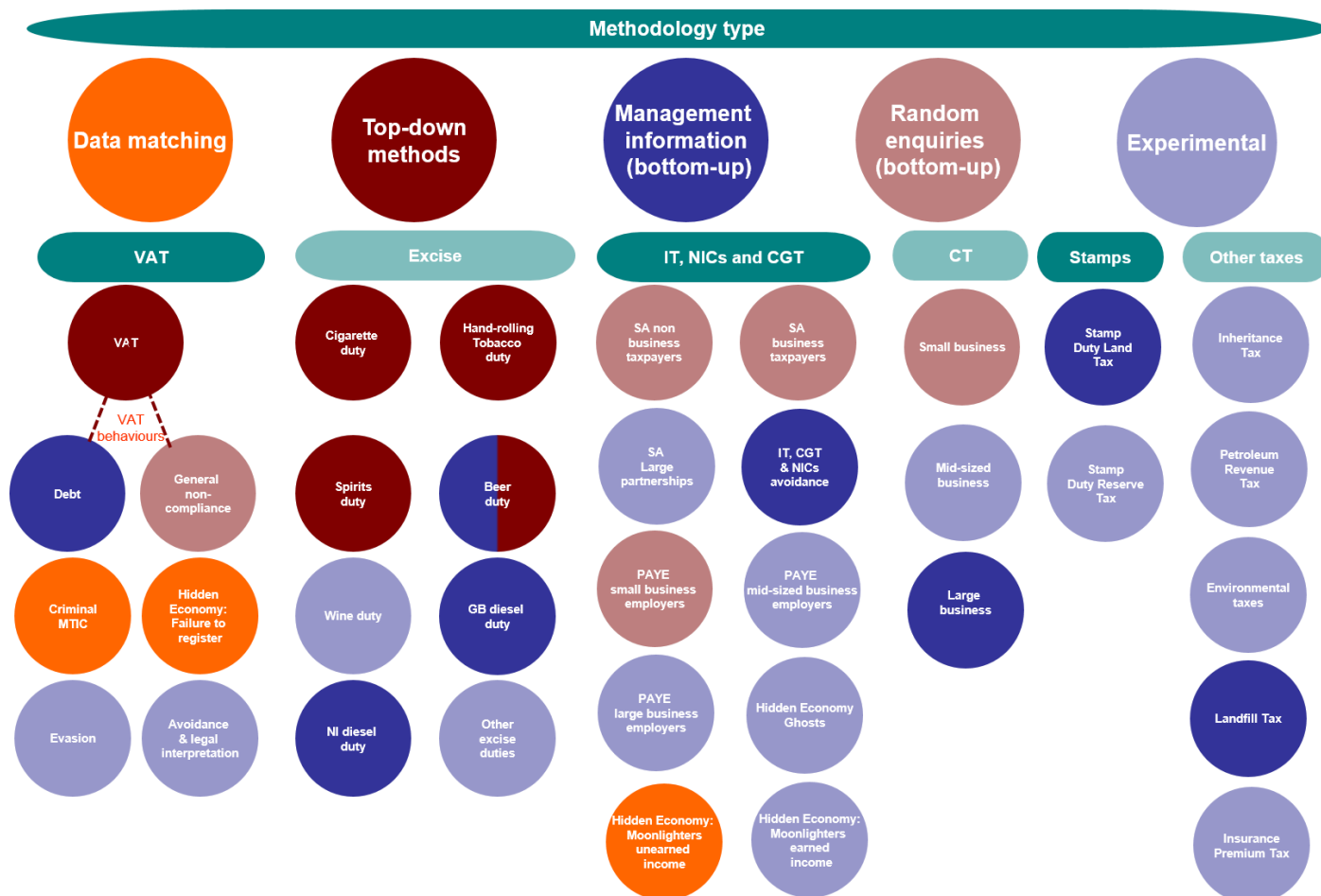
B.1   The figures presented in 'Measuring tax gaps 2018 edition' are subject to statistical uncertainty caused by two factors:

- sampling error – errors that arise because the estimates rely on information collected from a sample, rather than from the whole population; sampling error can lead to year-on-year fluctuations in the tax gap estimates that do not reflect true changes in the size of the tax gap

- bias or non-sampling error – systematic errors where the modelling assumptions or errors in the data lead to estimates that are consistently either too low or too high.

B.2   Where possible HMRC has estimated the likely impact of sampling errors by calculating statistical confidence intervals. These give margins of error within which the true value would be expected to lie in 95% of the time, if there were no systematic errors. They provide an indication of the extent to which changes in the estimates between years can be confidently interpreted as true changes. They do not take account of systematic errors that might lead the central estimate to be too low or too high over the whole series.

B.3   Systematic error is less straightforward to deal with, as it is not defined by statistical assessments that allow for easy interpretation. In order to give an indication of the effect of these biases HMRC presents the tax gaps for alcohol and tobacco as ranges. For beer and tobacco these are constructed as the range between upper and lower bounds, representing the degree of uncertainty associated with those systematic biases for which upper and lower bounds can be derived.

Value Added Tax

B.4   The VAT Total Theoretical Liability (VTTL) model and the top-down VAT gap derived from it are broad measures, subject to a degree of uncertainty. They are based on an analysis of survey and other data, and include a number of assumptions and adjustments which add both random and systematic variation to the estimates. There is a small element of forecasting in some of the spending data, which introduces further variation.

B.5   It is not possible to produce a precise confidence interval for the VAT revenue loss estimates. The VTTL estimate is constructed largely from Office for National Statistics (ONS) National Accounts data which is derived, in the main, from sample surveys and is thus subject to both sampling and non-sampling errors. The ONS does not publish error margins for the relevant input series and so it is not possible to construct an estimate of the impact of these errors on the VTTL.

B.6   The VAT gap is updated and revised as and when new data becomes available and new methodologies are developed.

Excise duties

Systematic biases

B.7   Systematic biases are explicitly considered for beer and tobacco products, with results presented as a range. These ranges are discussed in chapter E for beer and chapter F for tobacco products.

B.8   No account is presently made for systematic biases in the spirits and diesel estimates. Note that estimates for wine are illustrative from 2013-14, and take no account of systematic biases.

Random variation

B.9   While the upper and lower estimates for beer and tobacco will contain random variation, the resulting confidence intervals are not shown in this document as these estimates are used to represent the uncertainty around our central estimate.

B.10   For spirits and wine, an assessment of the effect of random variation is included using error margins. These are estimated by combining the random errors (where available) from all data sources used to calculate total consumption. These approximate to 95% confidence intervals, standard across statistical analyses.

B.11   For diesel, an assessment of the effect of random variation is included using the error margins resulting from the data used to estimate illicit consumption.

B.12   The central estimates for spirits and wine may not necessarily be half way between the upper and lower bounds as these bounds are confidence intervals, which may not be symmetric about the central estimate. As we do not have appropriate confidence intervals for the beer or tobacco tax gaps, the central estimate is calculated as the mid-point between the upper and lower estimates.

Direct taxes

Systematic biases

B.13   For direct tax estimates based on random enquiries, an adjustment is made to account for under-declarations that are not detected. HMRC continues to undertake analyses to define suitable ranges for other systematic biases in the direct tax estimates.

B.14   Direct tax gaps that rely on management information methods measure known components separately. There are also unknown factors that are not fully identified, leading to additional unmeasured losses.

Random variation

B.15   Direct tax estimates derived from random enquiries will be subject to random sampling errors. 95% confidence intervals have been calculated for these estimates using standard statistical techniques.

Chapter C: Tax Gap and Compliance Yield

C.1   Tax gap estimates are calculated net of compliance yield; that is, they reflect the tax gap remaining after HMRC compliance work.

C.2   In some cases we specifically use the cash collected measure of compliance yield in the tax gap calculation:

Tax Gap Component	Compliance Yield
Self Assessment (excluding large partnerships)	Deducted from gross tax gap; actual compliance yield series shown in table 4.1*.
Self Assessment (large partnerships)	Deducted from gross tax gap; actual compliance yield series shown in table 4.6*.
PAYE (small businesses)	Deducted from gross tax gap; actual compliance yield series shown in table 4.7*.
PAYE (mid-sized business)	Deducted from gross tax gap; actual compliance yield series shown in table 4.9*.
PAYE (large businesses)	Deducted from gross tax gap; actual compliance yield shown in table 4.10*.
Corporation Tax (large businesses)	Deducted from gross tax gap; compliance yield series shown in table 5.1*. This will represent both actual compliance yield (for closed cases) and estimates of compliance yield (for tax cases which are still under enquiry).
Corporation Tax (mid-sized businesses)	Deducted from gross tax gap; actual compliance yield series shown in table 5.2*.
Corporation Tax (small businesses)	Deducted from gross tax gap; actual compliance yield series shown in table 5.3*.
Diesel	Deducted from gross tax gap.
Landfill Tax	Deducted from gross tax gap.

*References to Tables in 'Measuring tax gaps 2018 edition'

C.3   In the following components of the tax gap we use an estimate of compliance yield as part of the calculation or do not take into account compliance yield:

Tax Gap Component	Compliance Yield
Avoidance (income tax, National Insurance Contributions and Capital Gains Tax)	Compliance yield is estimated by looking at the success of avoidance cases in a related area (large business) over time. Actual compliance yield is highly variable and doesn't relate to tax gap years.
Hidden Economy - ghosts	Does not currently take account of compliance yield.
Hidden Economy - moonlighters	Based on developing methodology which estimates the tax gap directly and does not currently take account of compliance yield.

C.4   In the remaining components of the tax gap we use a top-down method of calculation, looking at the difference between total theoretical liabilities and tax receipts. Although compliance yield is not explicitly included in these calculations it is reflected as part of tax receipts:

Tax Gap Component	Compliance Yield
VAT	Not explicitly used, but is reflected in receipts.
Tobacco	Not explicitly used, but is reflected in receipts.
Alcohol	Not explicitly used, but is reflected in receipts.
Stamp Duties	Not explicitly used, but is reflected in receipts.

C.5   HMRC publishes a detailed breakdown of compliance revenues within our Annual Report and Accounts. This differs in coverage and timing from the compliance information presented in 'Measuring tax gaps'.

Chapter D: VAT

VAT gap

General methodology

D.1   The VAT gap is measured by comparing the net VAT total theoretical liability (VTTL) with actual receipts – this is comparing the amount of VAT that HMRC expects to receive in the UK and the VAT HMRC actually receives. The difference between these amounts is known as the VAT gap. The VAT gap methodology uses a 'top-down' approach which involves:

- gathering data detailing the total amount of expenditure in the economy that is subject to VAT, primarily from the Office for National Statistics (ONS)

- applying the rate of VAT on the ONS expenditure data based on commodity breakdowns to derive the gross VTTL

- subtracting any legitimate refunds occurring through schemes and reliefs, to arrive at the net VTTL

- subtracting actual VAT receipts from the net VTTL

- leaving the residual element - the VAT gap, which includes, for example, error, evasion and debt.

D.2   The VTTL is the theoretical amount of VAT that would be collected in an ideal world. This means applying the rate of VAT on that expenditure where VAT should be payable, assuming that there is no fraud, avoidance, or losses due to error or non-compliance.

D.3   The VTTL includes irrecoverable VAT, which is the VAT paid on 'finally taxed expenditure' which cannot be reclaimed, for example by those not registered for VAT.

D.4   The expenditure data series used in the calculation are mainly constituents of National Accounts macroeconomic aggregates. All National Accounts data used to construct VTTL estimates is consistent with the ONS Blue Book.

Calculation of gross VTTL

D.5   The gross VTTL is calculated by multiplying the total amount of expenditure in the economy (also known as VAT-able expenditure) by the appropriate VAT rates.

D.6   For each of the expenditure sectors, the total expenditure is split according to the different VAT treatments; zero rated, standard rated, reduced rated and exempt. For the purposes of calculating the gross VTTL, only the standard and reduced rated expenditure are used.

D.7   The total VAT-able expenditure for each sector is combined together to represent an overall annual figure for the economy.

D.8   In order to derive the amount of VAT within the VAT-able expenditure, it is necessary to multiply the expenditure by the VAT fraction. The annual gross VTTL is thus calculated by multiplying the annual expenditure figure for the economy by the respective VAT fraction.

D.9   A number of streams of expenditure contribute to the tax base, with most VAT deriving from consumers' expenditure (that is, household consumption). The main expenditure categories that comprehensively cover VAT liabilities are:

- household consumption

- non-profit institutions serving households

- government capital and current expenditure

- VAT exempt sector capital and current expenditure

- housing capital expenditure.

Input tax adjustments

D.10　　Net VAT liability is the difference between VAT due on taxable supplies made by registered traders ('output tax'), and VAT recoverable by traders on supplies made to them ('input tax').

D.11　　VAT liability for the relevant categories can be estimated directly from ONS National Accounts data, with one exception - the VAT exempt sector. Businesses making outputs that are exempt from VAT are generally not permitted to reclaim all the VAT on inputs associated with their exempt outputs. In order to make an adjustment for this irrecoverable input tax, a separate HMRC survey is used to ascertain the proportion of purchases on which VAT cannot be reclaimed.

D.12　　A further adjustment is made for expenditure by businesses which are legitimately not registered for VAT and, as such, cannot recover their input tax. This adjustment uses a combination of data from the Department for Business, Energy and Industrial Strategy (BEIS) and HMRC information on the distribution of business turnover below the VAT threshold to estimate relevant expenditure.

D.13　　Finally, HMRC data and third party data sources are used in conjunction with National Accounts data to inform estimates of business expenditure on cars and entertainment, on which VAT is due.

D.14　　Because the calculation of irrecoverable input tax is complex, the level of uncertainty around input tax adjustments is larger than for the other elements.

Deductions

D.15　　The sum of the VAT liability arising from each of the expenditure categories listed in paragraph D.9 gives an estimate of the gross VTTL in each year. However there are a number of legitimate reasons why part of this theoretical VAT is not actually collected. These can be grouped into two broad categories:

- VAT refunds
- expenditure of traders legitimately not registered for VAT.

D.16　　VAT refunds are made primarily to government departments, NHS Trusts and regional health authorities for specified contracted out services acquired for non-business purposes. A number of other categories of expenditure cannot be separately identified in the overall VTTL calculation, for which VAT can be refunded. The value of these refunds is taken directly from audited HMRC accounts data.

D.17　　Traders who trade below the VAT threshold can legitimately exclude VAT on their sales. Expenditure on the output of these businesses will have been picked up in the total theoretical liability. To adjust for this, an estimate of relevant expenditure is made using a combination of BEIS data and HMRC information on the distribution of business turnover below the VAT threshold.

Net VAT receipts

D.18　　Figures for actual receipts of VAT are taken from HMRC's published Consolidated Fund figures. The receipts are adjusted to reflect timing effects within each financial year, before being used in the model.

VAT gap[1]

D.19　　Finally, subtracting the Net VAT Receipts from the Net VTTL derives the VAT gap. The percentage gap is further calculated by dividing the VAT gap by the Net VTTL. Receipts for the financial year (April to March) are compared with the total theoretical liability for the calendar year, assuming an average three month lag between an economic activity and the payment of the corresponding VAT to HMRC.

D.20　　The detailed calculations used to construct the estimated VTTL are continuously reviewed to identify improvements to the methodology. Also the National Accounts data used to construct the

---

[1] Calculations for VTTL and Net VTTL assume a three month lag between expenditure and actual VAT receipts. Hence, calendar year expenditure data equates to financial year receipts.

VTTL is subject to updates and revision by ONS throughout the year. This is part of the routine revisions to the ONS National Accounts data as final data become available.

Chapter E: Alcohol

Overview

E.1   The estimates of the illicit market for spirits, wine and the beer upper bound are produced using a top-down methodology. That is, the estimate is produced by first estimating total consumption, and then subtracting legitimate consumption, the residual being the illicit market:

> Illicit Market   =   Total Consumption   -   Legitimate Consumption

E.2   The above equation provides an estimate of the volume of goods supplied through the illicit market. This is then turned into an estimate of the proportion of the total market that is supplied through the illicit market, termed the illicit market share:

$$\text{Illicit Market Share} = \frac{\text{Illicit Market}}{\text{Total Consumption}} * 100$$

E.3   Revenue losses associated with the illicit market are then estimated by combining the illicit market share information with price data, excise duty and VAT rate information.

E.4   Details of the methodology for estimation of spirits, beer and wine tax gap are provided in the following sections.

Spirits central estimate

Estimating total consumption

E.5   Total consumption of spirits can be summarised as:

$$\text{Total Consumption} = \text{Total Consumption of UK Purchases} + \text{Cross-Border Shopping}$$

E.6   The consumption of spirits bought in the United Kingdom (UK) is estimated using the Family Spending Surveys (FSS) from the Office for National Statistics (ONS). This comprises of the Family Expenditure Survey for the period prior to 2001, the Expenditure and Food Survey (EFS) from 2001 to 2008 and the Living Costs and Food Survey (LCF) from 2008 to the present time. EFS and LCF estimates are weighted by the ONS to adjust for survey non-response.

E.7   The FSS only covers purchases within the UK. Because of this, total consumption is made up of the consumption of spirits bought in the UK plus the cross-border shopping.

Total consumption of UK purchases

E.8   The consumption of UK purchased goods in any given year is calculated using the following:

- estimates of household on-licence (consumed at the point of sale, for example, in a pub or restaurant) and off-licence (consumed off the premises, for example from a supermarket) expenditure on spirits from the FSS; per person expenditure is calculated using the average number of people in a household, also from FSS

- data on average prices provided by the ONS

- estimates of the UK adult population (18+) from the ONS

- uplift factors covering under-reporting. Uplift factors are calculated independently for on-licence and off-licence sectors as on-licence expenditure is more prone to under-reporting.

E.9 Total UK consumption is defined using this formula:

$$\text{Total UK Consumption} = \frac{\text{Average Household Consumption}}{\text{Average Number of Adults per Household}} * \text{UK Adult Population} * \text{Uplift Factor}$$

Family Spending Surveys

E.10 The average weekly expenditure on spirits for an average UK household is estimated using the FSS. Households participating in the surveys are asked to record their expenditure on alcohol under the relevant specific category of drink (that is wine, spirits, beer, etc.). There is an additional category for recording drinks purchased as part of a 'round' of drinks, which will be referred to as 'other drinks'.

E.11 Some of the 'other drinks' consumed will be spirits: the calculation for spirits consumption should therefore include a proportion of 'other drinks' consumption.

E.12 Following a change in interviewer guidance from the first quarter of 2014, whenever the value of an item recorded in 'other drinks' expenditure was greater than £10.00, interviewers were required to ask for further information, and if needed, re-allocate the item to the correct alcohol expenditure. This guidance change has enabled us to more accurately estimate the proportion of 'other drinks' expenditure that should have been allocated to beer and spirit expenditure historically. Revisions have been made across the historical time series to reflect this.

E.13 The average weekly expenditure per household is converted to the volume consumed by that household using the average price of spirits. This is then scaled up to an annual figure.

E.14 The average consumption of spirits per household is then converted to the average per person, by dividing by the average number of adults in a household. This is scaled up to the UK adult population.

E.15 Most under-age drinking is taken into account in the alcohol models. We assume that adults buy most of the alcohol consumed by minors. This under-age alcohol expenditure is therefore included in the adults' alcohol consumption, and is measured by the survey.

E.16 Due to the relatively small sample size in the FSS, the average weekly expenditure spirits is heavily influenced by extreme expenditure values in the data. Outliers in the data have been inspected and capped in keeping with the treatment of outliers elsewhere in the tax gap estimation.

Under-reporting uplift factor

E.17 The FSS data for alcohol are subject to under-reporting so an uplift factor is necessary to correct for this bias. This uplift factor is calculated by taking estimates of consumption from the FSS in the base year, and comparing these with independent estimates of total consumption.

E.18 To do this we take a year in which there is believed to be little or no illicit market, and use HMRC clearance data as a true indication of total consumption. In order to reduce sampling error, the uplift factor is derived by taking the average of three year's data: 1990-91; 1991-92; and 1992-93.

E.19 Separate uplift factors are calculated for on-license and off-licence markets, however the formula remains the same.

E.20 The uplift factors are defined as:

$$\text{Uplift Factor} = \frac{\text{Legitimate Consumption in 1990-91 to 1992-93}}{\text{Estimated Total Consumption in 1990-91 to 1992-93}}$$

with the appropriate legitimate and total consumption figures used.

Upper and lower confidence intervals

E.21   The confidence intervals indicate the potential size of chance fluctuations in the estimates due to sampling error. The estimates rely on a number of assumptions to combine the various data sources and produce an overall tax gap estimate. The confidence intervals do not take account of any systematic tendency to over or under-estimate the size of the tax gap that might arise from these assumptions. So the central estimate is best interpreted as an indicator of long-term trends in the illicit market rather than a precise estimate of the level or year-on-year changes.

E.22   The variation in the FSS is used to construct 95% confidence intervals around the central estimate. These are regarded as the upper and lower estimates, however they do not account for systematic errors in the model assumptions.

E.23   This leads to a central estimate defined by:

Total UK Consumption	=	Spirits Consumption	+	Proportion of 'Other' Drinks Consumption

Average Adult Spirits Consumption	*	UK Adult Population	*	Uplift Factor

Proportion of Average Adult 'Other' Drinks Consumption	*	UK Adult Population	*	Uplift Factor
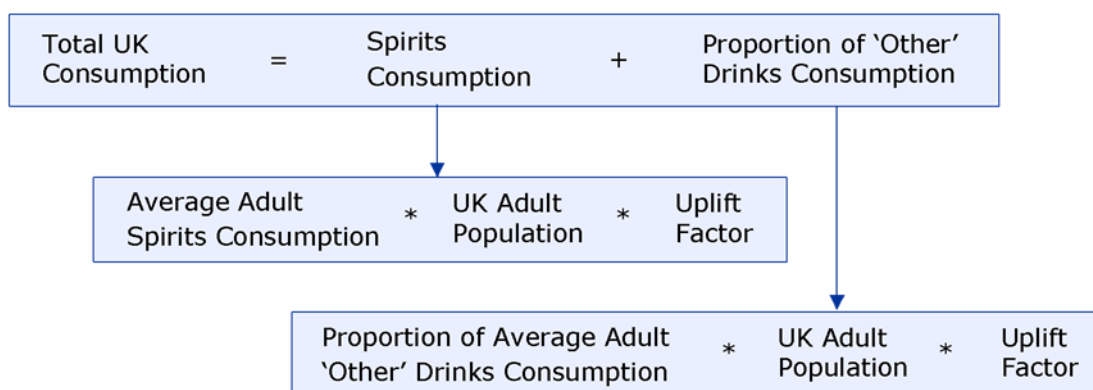
Cross-border shopping and duty free

E.24   Duty free is included in the cross-border shopping calculation. Estimates of consumption of goods purchased as cross-border shopping are based on figures produced from the International Passenger Survey (IPS). This provides estimates of the volume of spirits an average adult traveller brings into the country, separately for air and sea passengers. The IPS figures are weighted by the ONS, scaling up the survey data to represent the total cross-border shopping entering the UK.

E.25   An estimate of the volume of duty free spirits brought into the country is calculated in the same way, using passengers coming from outside the European Union (EU).

E.26   This estimate, however, does not cover sales made on-board ferries, so commercially provided data about deliveries of spirits to ferries are used to supplement the cross-border shopping estimate, and provide a complete figure.

E.27   Cross-border shopping is estimated as:

Cross-Border Shopping	=	Goods Bought Overseas	+	Goods Bought On-board Ferries	+	Duty Free
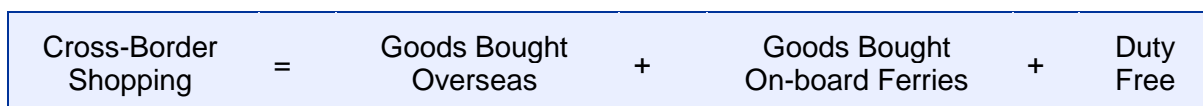
Estimating legitimate consumption

E.28   Estimates of legitimate consumption have two elements; UK duty paid consumption and cross-border shopping:

Legitimate Consumption	=	UK Duty Paid Consumption	+	Cross-Border Shopping
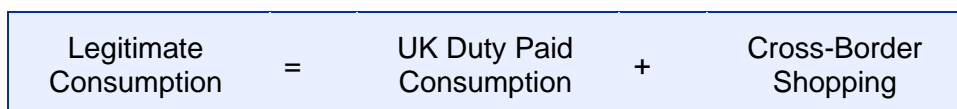
E.29   Estimates of UK duty paid consumption are taken directly from returns to HMRC of the volumes of spirits on which duty has been paid. The volumes of ready-to-drink products have been removed from spirits clearances in order to obtain figures for spirits only.

E.30   Cross-border shopping is calculated in the same way as for total consumption:

Cross-Border Shopping	=	Goods Bought Overseas	+	Goods Bought On-board Ferries	+	Duty Free

Estimating the illicit market

E.31   Combining the two formulae for total consumption and legitimate consumption gives the formula for the illicit market:

$$\text{Illicit Market} = \text{Total Consumption} - \text{Legitimate Consumption}$$

$$\text{Legitimate Consumption} = \text{UK Duty Paid Consumption} + \text{Cross-Border Shopping}$$

$$\text{Total Consumption} = \text{Total Consumption of UK Purchases} + \text{Cross-Border Shopping}$$

Conversion to monetary losses

E.32   Revenue losses associated with the illicit market are then estimated by combining the illicit market share information with price data and duty and VAT rate information, using this formula:

$$\text{Financial losses} = (\text{Illicit Volume} * \text{Spirits Duty}) + (\text{Illicit Volume} * \text{Average Price} * \text{VAT Fraction})$$

E.33   Data on average spirits prices is derived from data provided by the ONS. The prices used in the model are weighted across on and off trade and for different types of spirits.

E.34   The VAT fraction is the portion of the retail price that is VAT – for example, a 20% VAT rate is equivalent to a 1/6 VAT fraction. VAT fractions are calculated annually to capture changes in the VAT rate. This method assumes that VAT is also lost on all purchases. As, in some cases, the final illicit product is sold in legitimate outlets this may not always be the case, and this will be an overestimate of losses.

E.35   Spirits duty is converted into bulk duty liabilities based on the assumption that spirit's strength is constant at 38%.

Summary of methodology

E.36   A summary of the calculation of the illicit market central estimate is:

Illicit Market   =   Total Consumption   -   Legitimate Consumption

UK Duty Paid Consumption   +   Cross-Border Shopping

Average Adult Spirits Consumption   *   UK Adult Population   *   Uplift Factor

+

Proportion of Average Adult 'Other' Drinks Consumption   *   UK Adult Population   *   Uplift Factor

+

Cross-Border Shopping
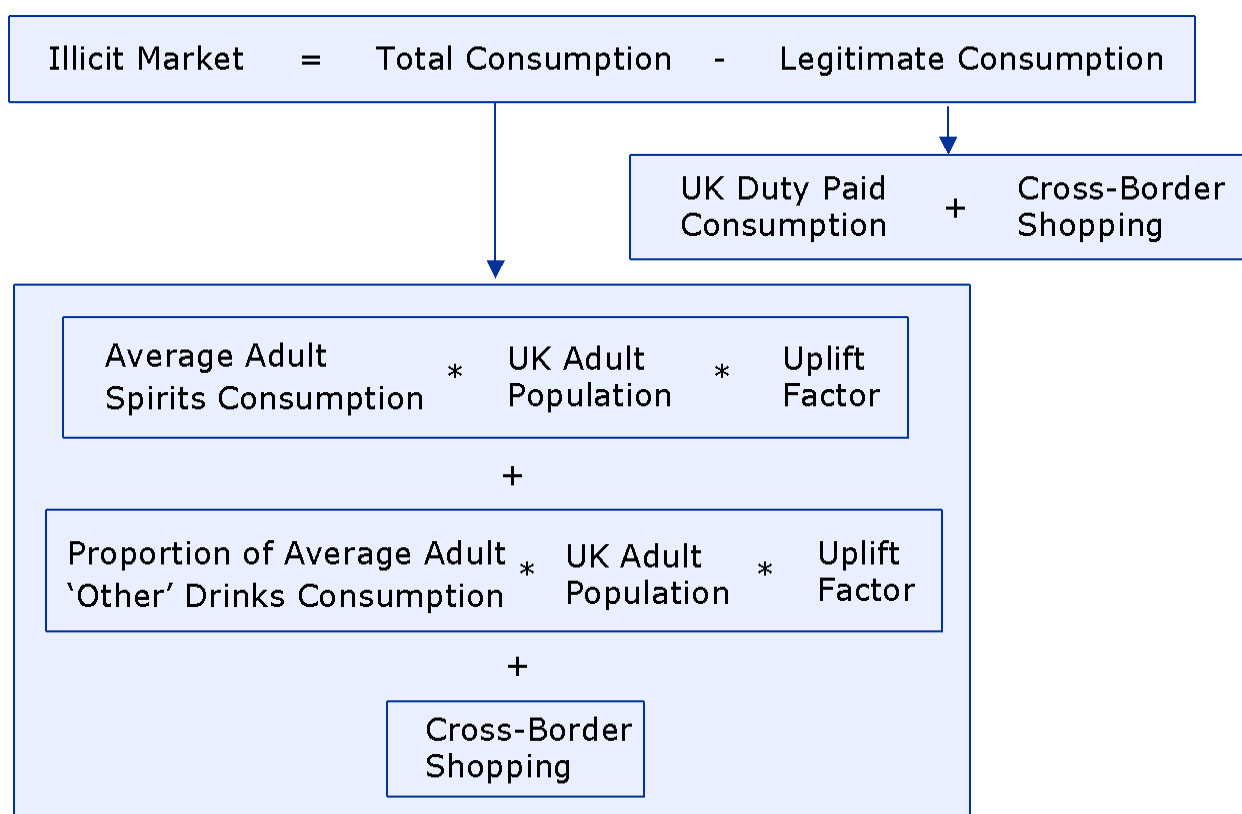
Beer upper estimate

Methodology

E.37   The estimate of the illicit market for beer is produced using a top-down methodology as described in paragraphs E.1 to E.4.

E.38   Details of the estimation of total consumption and of legitimate consumption are provided in the subsequent sections.

Total consumption

E.39   The consumption of beer purchased in the UK for each year is estimated using the FSS. FSS estimates are weighted by the ONS to adjust for the survey non-response.

E.40   The FSS only covers purchases within the UK. Because of this, total consumption is made up of consumption of beer bought in the UK (from FSS) plus the cross-border shopping of beer.

UK purchases

E.41   The consumption of UK purchased goods in any given year is calculated using the following:

- estimates of households' on-licence (consumed at the point of sale – for example in a pub or restaurant) and off-licence (consumed off the premises – for example from a supermarket) expenditure on beer from FSS. HMRC converts these into volumes consumed using data on average prices provided by ONS. Due to the relatively small sample size in the FSS, specifically the LCF survey, the average weekly expenditure on beer is heavily influenced by extreme expenditure values in the data. Outliers in the data have been inspected and capped in keeping with the treatment of outliers elsewhere in the tax gap estimation

- a proportion of 'other' drinks consumption, which is added to beer consumption

- estimates of draught and packaged beer consumption from on-licence and off-licence consumption using data published in the British Beer and Pub Association (BBPA) handbook

- estimates of the UK adult population (18+) from the ONS

- estimates of the number of adults per household in the UK, from FSS

- an uplift factor covering under-reporting. Uplift factors for both draught and packaged beer are calculated.

E.42   UK purchases are defined as:

$$\text{UK Purchases} = \text{Average Adult Consumption} * \text{UK Adult Population} * \text{Uplift Factor}$$

$$\text{Average Adult Consumption} = \frac{\text{Average Household Consumption of Beer} + \text{Proportion of 'Other Drinks'}}{\text{Average Number of Adults per Household}}$$

Other drinks

E.43   Some of the 'other' drinks consumed will be beer, so the calculation for beer consumption should include a proportion of 'other' drinks consumption. Details of the 'other' drinks methodology are in paragraphs E.10 to E.12.

Uplift factor

E.44   The FSS data is subject to under-reporting so an uplift factor is necessary to correct for this bias. The base year uplift factor is calculated by taking estimates of consumption from the FSS in a base year, and comparing these with independent estimates of total consumption. To do this we take a year in which there is believed to be little or no illicit market, and use HMRC clearance data as a true indication of total consumption. In order to reduce sampling error, the uplift factor is calculated by taking an average of three year's data: 1990-91, 1991-92 and 1992-93.

E.45   To calculate uplift factors for draught and packaged beer, FSS data is split between on-licence and off-licence markets and then into draught and packaged beer. This uses market shares estimated from ONS and BBPA data.

E.46   The base year uplift factors are defined as:

$$\text{Base Year Uplift Factor} = \frac{\text{Legitimate Consumption in 1990-91 to 1992-93}}{\text{Estimated Total Consumption in 1990-91 to 1992-93}}$$

E.47   An additional uplift for packaged beer is calculated which varies year-on-year. This is based on the assumption that there is no or negligible illicit market in draught beer, so draught beer consumption should be equal to draught clearances in every year. This assumption gives an uplift for draught beer in each year. The draught beer uplift and base year uplifts are combined to compute the packaged beer uplift, using this formula:

$$\text{Packaged Uplift} = \text{Draught uplift} * \frac{\text{Base Year Packaged Uplift}}{\text{Base Year Draught Uplift}}$$

Cross-border shopping and duty free

E.48   Estimates of consumption of goods purchased as cross-border shopping are based on figures produced from IPS. This provides estimates of the volume of beer brought back by UK travellers. The IPS figures are weighted by the ONS to represent the total cross-border shopping entering the UK.

E.49   This estimate does not cover sales made on-board ferries, so commercially provided data for deliveries of beer to ferries are used to supplement the cross-border shopping estimate, and provide a complete figure.

E.50   An estimate of the volume of duty free beer brought into the country is calculated in the same way, using passengers coming from outside the EU.

E.51   Cross-border shopping is estimated as:

Cross-Border Shopping	=	Goods Bought Overseas	+	Goods Bought On-board Ferries	+	Duty Free

E.52   Total consumption can be summarised as:

Total Consumption	=	Total Consumption of UK Purchases	+	Cross-Border Shopping

Legitimate consumption

E.53   Estimates of legitimate consumption have two elements:

- UK duty paid consumption
- cross-border shopping.

UK duty paid consumption

E.54   Estimates of UK duty paid consumption are taken directly from returns to HMRC (clearances data) of the volumes of beer on which duty has been paid, along with the actual amounts of money.

Cross-border shopping

E.55   This is calculated in the same way as for total consumption:

Cross-Border Shopping	=	Goods Bought Overseas	+	Goods Bought On-board Ferries	+	Duty Free

E.56   Legitimate consumption can be summarised as:

Legitimate Consumption	=	UK Duty Paid Consumption	+	Cross-Border Shopping

Illicit market

E.57   Subtracting legitimate consumption from total consumption will give an estimate of the illicit market. This can be written as:

$$\text{Illicit Market} = \text{Total Consumption} - \text{Legitimate Consumption}$$

UK Purchases + Cross-Border Shopping	UK Clearances + Cross-Border Shopping
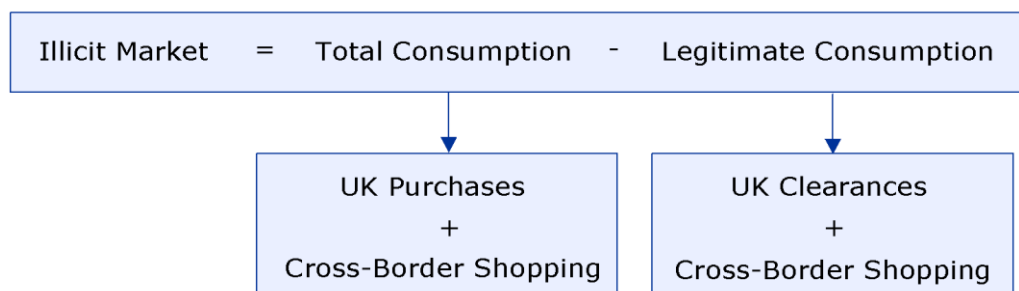
Conversion to monetary losses

E.58   Revenue losses associated with the illicit market are then estimated by combining the illicit market share information with price data, duty and VAT rate information, using the following formula:

$$\text{Financial losses} = (\text{Illicit Volume} * \text{Spirits Duty}) + (\text{Illicit Volume} * \text{Average Price} * \text{VAT Fraction})$$

E.59   Where:

- average beer prices are derived from the data provided by the ONS

- the VAT fraction is the portion of the retail price which is VAT – for example, a 20% VAT rate is equivalent to a 1/6 VAT fraction. VAT fractions are calculated annually to capture changes in the VAT rate.

E.60   This method of converting volume into monetary losses assumes that VAT is also lost on all illicit beer. As, in some cases, the final illicit product is sold in legitimate outlets this may not always be the case, and this will be an overestimate of losses.

Summary of methodology

E.61   A summary of the calculation of the illicit market upper estimate is:

$$\text{Illicit Market} = \text{Total Consumption} - \text{Legitimate Consumption}$$

UK Purchases + Cross-Border Shopping	UK Clearances + Cross-Border Shopping

HMRC Clearance Data

Goods Bought Overseas + Goods Bought On-board Ferries

Average Adult Consumption of Beer * UK Adult Population * Uplift Factor

$$\frac{\text{Average Household Consumption of Beer} + \text{A Proportion of 'Other Drinks'}}{\text{Average Number of Adults per Household}}$$

Beer lower estimate

Overview

E.62   The beer tax gap lower estimate is produced using a bottom-up methodology. This means estimates of the illicit market are made directly, by estimating the fraud components that make up the illicit market. The following types of illicit beer are included in the lower estimate:

- diversion of UK-produced beer

- drawback fraud.

E.63   Some of this illicit beer is recovered through HMRC compliance activity, so this is subtracted to give the net tax gap. The tax gap estimate is defined by:

Beer Illicit Market Lower Estimate	=	Diversion of UK Produced Beer	+	Drawback Fraud	-	Seizures of Illicit Beer

E.64   A number of beer frauds are not included in this methodology as we are currently unable to estimate them. This is one of the reasons it is a lower bounding estimate. These include:

- smuggled beer

- diversion of foreign produced beer

- counterfeit beer.

E.65   There may also be other types of fraud which we do not know about; again, these are not covered by the estimate.

Diversion of UK-produced beer

E.66   Diversion fraud occurs when beer is moved in duty suspense to the EU and is subsequently diverted back into the UK under the cover of false documentation. The taxes are not declared on the beer and the illicit product enters the UK market.

E.67   We estimate that diversion fraud is equal to the amount of beer moved in duty suspense from the UK to certain EU member states, minus legitimate demand for UK branded beer in those countries. That is, we assume that any UK beer which is not feeding demand abroad will be diverted back to the UK illicit market:

Diversion of UK Produced Beer	=	Duty Suspended Beer Moved to Selected EU Countries	-	Legitimate Demand in Selected EU Countries

E.68   The total amount of beer moved in duty suspense from the UK to the EU includes dispatches from both excise warehouses and brewers. Dispatches from excise warehouses are taken directly from Excise Warehouse Returns (W1 form). Dispatches from brewers are estimated using data from Beer Duty Returns (EX46 form):

Total Beer Dispatches	=	Dispatches from Warehouses	+	Dispatches from Brewers

E.69   Brewers return data is used for dispatches (movements to EU countries) and exports (movements to non-EU countries) and it cannot be disaggregated. So, to estimate dispatches from brewers, we subtract an estimate of exports from brewers.

E.70   Exports from brewers are estimated as total exports, from Customs Handling of Import and Export Freight (CHIEF), minus exports from Excise Warehouse Returns (W1 form).

E.71   To preserve the lower bounding nature of this estimate, we only include dispatches to certain EU countries. These countries have been selected based on a number of factors, including: proximity to the UK; the differential in price; operational indications of risk and patterns of supply.

$$\text{Dispatches from Brewers} = \text{Dispatches and Exports from Brewers} - \text{Exports from Brewers}$$

$$\text{Total Exports} - \text{Exports from Warehouses}$$
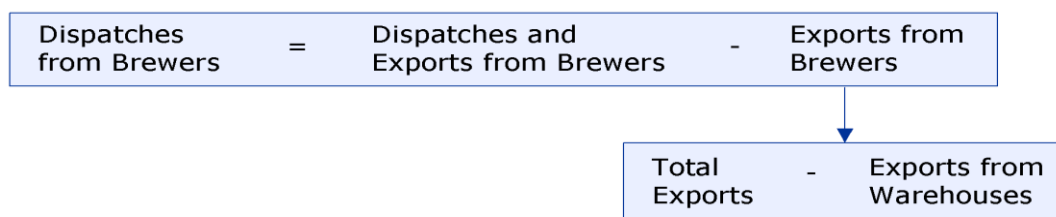
E.72   The estimate of beer dispatches, described in E.69 and E.72, cannot be broken down to the recipient country. So we use an alternative data source, UK trade data, which does include a breakdown by country. The proportion of beer dispatched to the selected EU countries is taken from UK trade data and applied to the estimated total dispatches to produce an estimate for dispatches to these selected EU countries.

E.73   UK trade data is not used to directly estimate dispatches to these countries as it does not include certain types of movements. More detail is provided on this in section E.87 of this annex.

E.74   To summarise:

$$\text{Duty Suspended Beer Moved to Selected EU Countries} = \text{Total Dispatches to EU Countries} \ast \text{Percentage of Dispatches going to Selected EU Countries}$$

$$\text{Dispatches from Warehouses} + \text{Dispatches from Brewers}$$

$$\text{Dispatches and Exports from Brewers} - \text{Exports from Brewers}$$

$$\text{Total Exports} - \text{Exports from Warehouses}$$

Drawback fraud

E.75   Drawback fraud occurs when goods are moved to the EU and the duty is reclaimed via drawback. Duty is then paid at the lower rate in the destination country and the goods are illicitly returned to the UK.
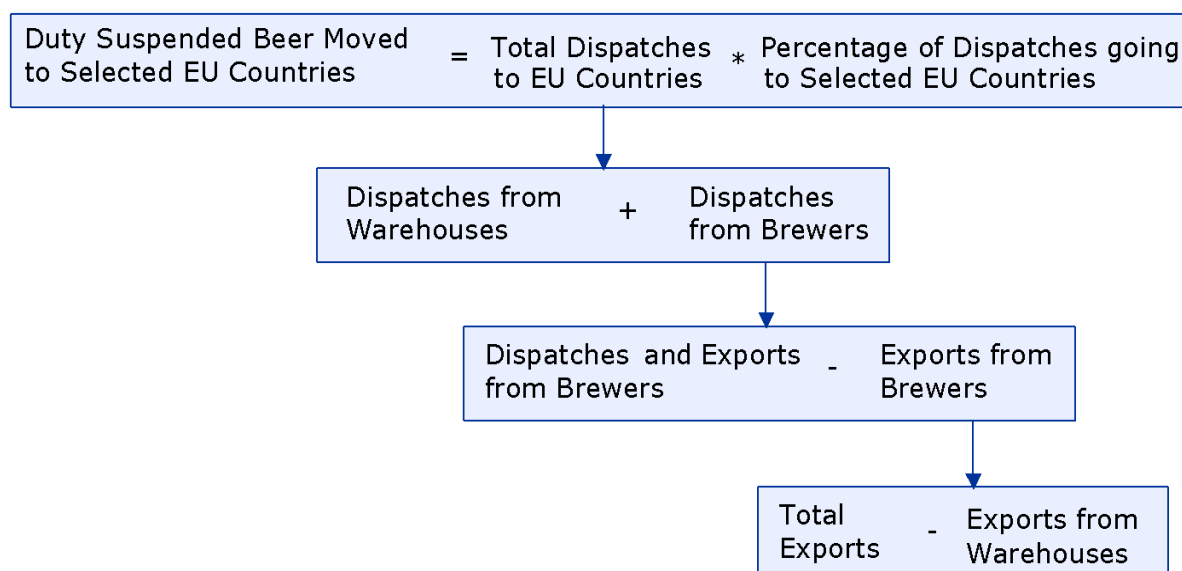
E.76   To estimate drawback fraud, we estimate the volume of beer corresponding to certain drawback claims, then subtract the legitimate demand for beer in the selected destination countries:

$$\text{Drawback Fraud} = \text{Selected Drawback Claims, Converted to Volume of Beer} - \text{Legitimate Demand in Selected EU Countries}$$

E.77   To preserve the lower bounding nature of this estimate, we only include drawback if it is claimed for dispatch by a business not part of HMRC Large Business (LB). The value of these drawback claims are converted to volumes of beer by dividing by the average duty rate for beer.

E.78   The volume is then adjusted using the proportion of dispatches going to the selected EU countries. This gives an estimate of the amount of beer going to the selected countries with drawback claimed by small and medium sized enterprises:

$$\text{Selected Drawback} = \frac{\text{Value of Drawback Claims (non LB, for Dispatch)}}{\text{Average Beer Duty Rate}} * \text{Percentage of Dispatches going Selected EU Countries}$$

Legitimate demand in selected EU countries

E.79   Some of the beer moved to the selected EU countries will be supplying legitimate demand within those countries, rather than being diverted to the UK illicit market. We make one overall estimate of legitimate demand in the selected EU countries, and subtract it from the sum of selected beer dispatches and selected beer for drawback.

E.80   We have purposely overestimated legitimate demand as this produces an underestimate of the illicit market, in order to maintain the lower bounding nature of the tax gap estimate.

E.81   The estimate of legitimate demand in other countries includes:

- cross-border shopping bought by UK residents

- legitimate consumption abroad, which may include: consumption by UK expatriates; consumption by UK residents on vacation or while working abroad; consumption by foreign nationals; beer in transit to other countries.

$$\text{Legitimate Demand in Selected EU Countries} = \text{Cross-Border Shopping by UK Residents} + \text{Legitimate Consumption Abroad}$$

E.82   Cross-border shopping is estimated using data from the IPS. More detail is provided in paragraph E.95. Only passengers from the selected EU countries are included.

Legitimate consumption of UK produced beer abroad

E.83   We could not find reliable data regarding legitimate consumption of UK produced beer abroad. So, we estimate it based on the assumption that in a certain year when the illicit market upper estimate was low, there was negligible illicit activity so all dispatches to the selected EU countries were consumed legitimately. This is likely to provide an overestimate of legitimate consumption abroad, as there would likely be some level of fraud in these years. This supports the methodology being a lower estimate of the tax gap.

E.84   For stability, an average of two years is used: 2000-01 and 2001-02. For simplicity, we will refer to these two years as the 'base year'.

E.85   Brewers return data is not available for years prior to 2007. Consequently we use an alternative data source, UK trade data, to estimate dispatches in the base year.

E.86   In the base year we assume that all dispatches supply either cross-border shopping by UK residents or legitimate consumption abroad. We subtract an estimate of cross-border shopping in the base year from dispatches in the base year; the remainder is assumed to be legitimate consumption abroad:

$$\text{Legitimate Consumption Abroad} = \text{Dispatches to Selected EU Countries in Base Year} - \text{Cross-Border Shopping from Selected EU Countries in Base Year}$$

E.87   We believe that UK trade data may underestimate beer dispatches in the base year as it does not record certain types of beer movement. These include:
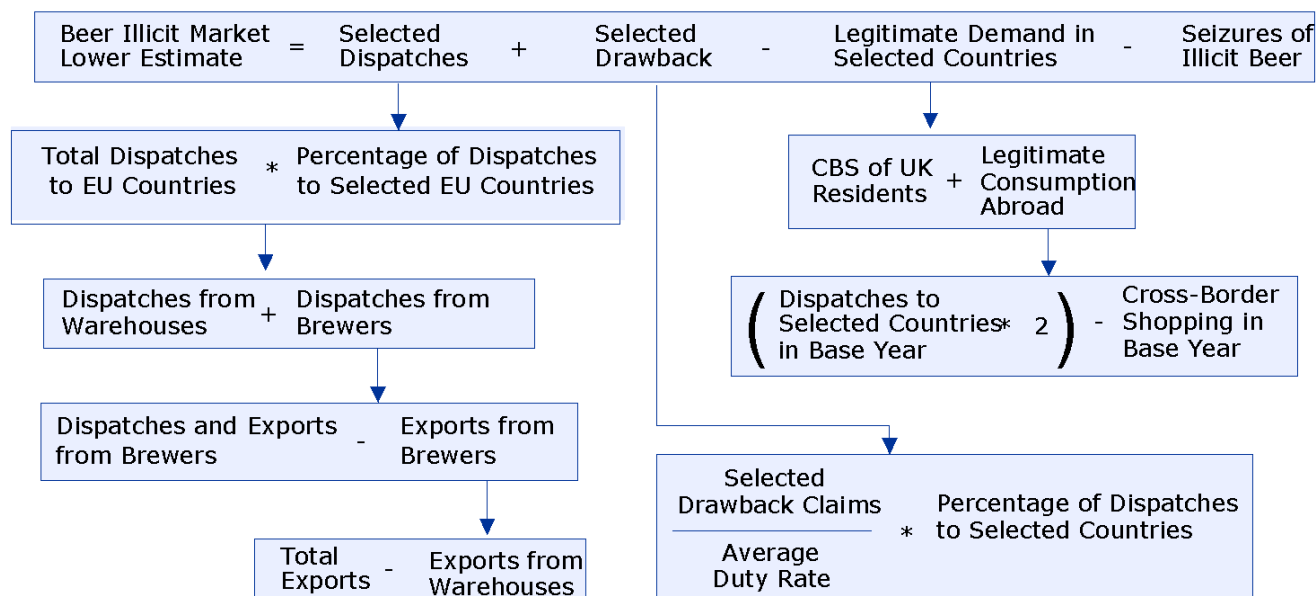
- goods in transit

- deliveries to embassies
- deliveries to Navy, Army and Air Force Institutes (NAAFI).

E.88 Additionally, as the threshold for recording goods on UK trade data is relatively high in beer terms, beer may have a higher proportion of small traders than other commodities. This may mean the standard adjustment applied to UK trade data to account for small traders may be too low for beer.

E.89 To account for these concerns we uplift the UK trade data. There is very little evidence to indicate the actual level of under-reporting. Comparison with our calculated dispatches in later years led us to apply a factor of two. Again, the high level of this adjustment may result in this being an overestimate, but this is in keeping with the lower bounding methodology for the tax gap.

Illicit market lower estimate

E.90 In summary, the illicit market is estimated as:

$$\text{Beer Illicit Market Lower Estimate} = \text{Selected Dispatches} + \text{Selected Drawback} - \text{Legitimate Demand in Selected Countries} - \text{Seizures of Illicit Beer}$$

$$\text{Selected Dispatches} = \text{Total Dispatches to EU Countries} * \text{Percentage of Dispatches to Selected EU Countries}$$

$$\text{Total Dispatches to EU Countries} = \text{Dispatches from Warehouses} + \text{Dispatches from Brewers}$$

$$\text{Dispatches from Brewers} = \text{Dispatches and Exports from Brewers} - \text{Exports from Brewers}$$

$$\text{Exports from Brewers} = \text{Total Exports} - \text{Exports from Warehouses}$$

$$\text{Legitimate Demand in Selected Countries} = \text{CBS of UK Residents} + \text{Legitimate Consumption Abroad}$$

$$\text{CBS of UK Residents} = (\text{Dispatches to Selected Countries in Base Year} * 2) - \text{Cross-Border Shopping in Base Year}$$

$$\text{Selected Drawback} = \frac{\text{Selected Drawback Claims}}{\text{Average Duty Rate}} * \text{Percentage of Dispatches to Selected Countries}$$

Illicit market share

E.91 The direct estimate of the volume of illicit beer is converted into an estimate of the proportion of the total market that is supplied through the illicit market termed the illicit market share:

$$\text{Illicit Market Share} = \frac{\text{Illicit Market}}{\text{Total UK Consumption}} * 100$$

Total UK consumption

E.92 Total UK consumption is calculated as the sum of legitimate UK consumption and the illicit market:

$$\text{Total UK Consumption} = \text{Legitimate UK Consumption} + \text{Illicit Market}$$

Legitimate UK consumption

E.93 Estimates of legitimate consumption include:

- UK duty paid consumption
- cross-border shopping and duty free.

UK duty paid consumption

E.94 Estimates of UK duty paid consumption are taken directly from returns to HMRC (clearances data) of the volumes of beer on which duty has been paid, along with the actual amounts of money.

Cross-border shopping and duty free

E.95   Estimates of the consumption of goods purchased as cross-border shopping are based on figures produced from the IPS. This provides estimates of the volume of beer brought back by UK travellers. The IPS figures are weighted by the ONS to represent the total cross-border shopping entering the UK.

E.96   This estimate does not cover sales made on-board ferries, so commercially provided data for deliveries of beer to ferries are used to supplement the cross-border shopping estimate, and provide a complete figure.

E.97   An estimate of the volume of duty free beer brought into the country is calculated in the same way, using passengers coming from outside the EU.

E.98   Cross-border shopping and duty free are estimated as:

Cross-Border Shopping and Duty Free	=	Goods Bought Overseas	+	Goods Bought On-board Ferries	+	Duty Free

Implied mid-point estimate

E.99   The implied mid-point estimate is calculated as the average of the upper and lower estimates. It is only intended as an indicator of long-term trend – the true tax gap could lie anywhere within the bounds.

E.100   The upper and lower estimates should be interpreted as indicators of long-term trend, rather than precise estimates of the level or of year-on-year changes. The bounds do not take account of any systematic tendency to over- or under-estimate the size of the tax gap that might arise from the modelling assumptions.

Wine central estimate

E.101   The wine tax gap previously used a top-down methodology that estimated the tax gap for wine, sparkling wines and made wines. For 2013-14 to 2016-17, an estimate using this methodology was not possible due to changes in data sources. An illustrative estimate has been provided for these years, based on the average of the estimates for 2011-12 and 2012-13. The methodology described below relates to estimates prior to 2013-14.

E.102   The calculation of the wine tax gap was similar to that of the spirit tax gap, with the only differences listed below:

- wine prices are taken as per litre weighted average of wine, sparkling wine and made wine

- wine clearances are the sum of wine, sparkling wine and made wine

- off-licence wine expenditure is the sum of wine, sparkling wine and made wine as recorded in the FSS

- off-licence wine consumption is calculated in the same way as for the spirits tax gap (see paragraphs E.8 to E.23).

E.103   On-licence consumption was estimated from off-licence consumption. We could estimate from commercially available data what the proportion each year of on-licence trade was relative to off-licence trade. We used this factor to estimate the on-licence volume:

On-licence Consumption = Off-licence Consumption * Commercial Data Weighting

24

E.104   The commercial data weighting was calculated as:

$$\text{Commercial Data Weighting} = \frac{\text{On-licence Proportion}}{\text{Off-licence Proportion}}$$

E.105   Total consumption was calculated as:

$$\text{Total Consumption} = (\text{Off-licence Consumption} * \text{Commercial Data Weighting}) + \text{Off-licence Consumption}$$

E.106   The uplift for wine was estimated using data from the year 2000.

E.107   Using a random Her Majesty's Customs and Excise passenger and freight survey in the year 2000, we were able to estimate the illicit market for the year 2000. This, in addition to cross-border shopping and clearances, allowed us to estimate the total wine consumption for 2000 as:

$$\text{Total Consumption}_{2000} = \text{Clearances}_{2000} + \text{Cross-Border Shopping}_{2000} + \text{Illicit}$$

E.108   Using the total consumption for 2000 and the FSS driven estimates of total consumption allowed us to infer an under reporting uplift as follows:

$$\text{Uplift Factor} = \frac{\text{Clearances} + \text{Cross-Border Shopping} + \text{Tax Gap}}{\text{Off-licence Consumption} + (\text{Off-licence Consumption} * \text{Commercial Data weighting})}$$

Summary of methodology

E.109   A summary of the calculation for the illicit market is shown below:

Illicit Market =   Total Consumption   –   Legitimate Consumption

UK Purchases + Cross-Border Shopping

UK Clearances + Cross-Border Shopping

HMRC Clearance Data

Goods Bought Overseas + Goods Bought On-board Ferries

On-licence Purchases + Off-licence Purchases

Average Adult Consumption of Wine * UK Adult Population * Uplift Factor

Off-licence Purchases* Commercial Data Weighting

$$\frac{\text{Average Household Consumption of Wine} + \text{a proportion of 'Other Drinks'}}{\text{Average Number of Adults per Household}}$$

$$\text{Uplift Factor} = \frac{\text{Clearances} + \text{Cross-Border Shopping} + \text{Tax Gap}}{\text{Off-licence Consumption} + (\text{Off-licence Consumption} * \text{Commercial Data Weighting})}$$

Chapter F: Tobacco

Overview

F.1   The estimate of the illicit market for tobacco is produced using a top-down methodology. That is, first we estimate total consumption, and then we subtract legitimate consumption. The residual is estimated to be the illicit market.

> Illicit Market   =   Total Consumption   –   Legitimate Consumption

F.2   The above equation provides an estimate of the volume of goods supplied through the illicit market. This is then turned into an estimate of the proportion of the total market that is supplied through the illicit market.

$$\text{Illicit Market Share} \quad = \quad \frac{\text{Illicit Market}}{\text{Total Consumption}} \quad * \; 100$$

F.3   Revenue losses associated with the illicit market are then estimated by combining the illicit market share information with price data, excise duty and VAT rate information.

F.4   Details of the methodology for both cigarettes and hand-rolling tobacco are provided in the following sections.

Cigarette estimates

Methodology

F.5   The estimate of the illicit market for cigarettes is produced using a top-down methodology as described in paragraphs F.1 to F.4.

F.6   Details of the estimation of total consumption and of legitimate consumption are provided in the subsequent sections.

Total consumption

F.7   The total consumption in any given year is calculated using the following:

- estimates of prevalence (proportion of the population that smokes cigarettes) from the General Lifestyle Survey (GLF), the Opinions and Lifestyle Survey (OPN) and Health Survey for England (HSE)

- estimates of cigarette consumption per smoker from GLF, OPN and HSE

- estimates of the adult population (16+) from the Office for National Statistics (ONS)

- an uplift factor covering under-reporting.

F.8   The estimate of total UK consumption of cigarettes for each year is based on the estimates of cigarette smoking prevalence and consumption per smoker for declared and undeclared smokers.

F.9   In general, most smokers admit that they smoke and we use survey data to estimate prevalence and consumption per smoker. Up to the end of 2011, the prevalence and consumption per smoker of these declared smokers was obtained from the GLF, formerly known as the General Household Survey (GHS). The GLF ceased at the end of 2011, so from 2012 onwards this was estimated using the OPN. This is a similar and well established survey, and has been used in calculations of the tax gap for hand-rolling tobacco in previous years.

F.10   The GLF and OPN are both administered by the ONS, but the main difference is the presence of dual smokers. These are respondents in the OPN survey that said they smoked both cigarettes and hand-rolling tobacco. This was not an option for respondents in the GLF.

F.11   There are some smokers who, for whatever reason, do not admit that they smoke. The prevalence and consumption of 'Undeclared Smokers' in the non-smoking population can be obtained from the HSE.

F.12   From 2013, the methodology for calculating the upper and lower bounds for cigarette consumption was improved. Previously the upper bound was based on consumption in the base year, with the lower bound based on the consumption in the reference year. As time passed, and as smoking habits changed, this assumption gave increasingly extreme and unrealistic bounds.

F.13   The new method for calculating the bounds is described below and is based on the same assumptions used for producing the hand-rolling tobacco estimates. The methodology and assumptions for the cigarette model has been retrospectively applied to all years back to 2000-01. As a result, the figures are produced on a consistent basis and trends in the illicit market can be directly compared.

Uplift factor

F.14   The survey data for tobacco consumption is subject to under-reporting due to the self-reporting nature of the surveys. An uplift factor is necessary to correct for the bias. This uplift factor is calculated by taking estimates of total consumption from the GLF in a base year, and comparing with consumption based on actual clearances to HMRC and an estimate of legitimately purchased cigarettes from abroad. We take a base year in which the illicit market can be measured using another source of data, and is believed to be small. For cigarettes, the year 1996-97 was used as this is the earliest year where we have results from the Cross-Channel Smuggling Survey, and hence an estimate of the tax gap.

F.15   The uplift factor is defined as:

$$\text{Uplift Factor} = \frac{\text{Legitimate Consumption in 1996-97 (based on HMRC clearances and estimate of duty free/cross-border shopping)}}{\text{Total Consumption in 1996-97 (based on self-reported consumption from GLF survey)}}$$

F.16   Due to uncertainties in the data sets used to create these estimates, it is not possible, with sufficient accuracy, to produce a single point estimate of total consumption. So, the exercise is undertaken twice – once to produce an upper bound for total consumption, and once to produce a lower bound. This allows us to produce a range for total consumption that takes account of the uncertainty in the underlying data.

Upper and lower bounds for total consumption

F.17   The one difference between the upper and lower bound calculations is the treatment of dual smokers. In the upper bound calculation the majority of the dual smokers are considered to be cigarette smokers. In the lower bound estimate, we assume that the majority of them smoke hand-rolling tobacco. This is explained further in the table below and in the following section.

F.18   Upper and lower bound assumptions:

OPN Survey Options	Allocation of total tobacco consumption to cigarette consumption	
	Upper bound assumption	Lower bound assumption
Cigarettes Only	100%	100%
Dual Smokers: Cigarettes and hand-rolling tobacco, but mainly cigarettes	99%	51%
Dual Smokers: Cigarettes and hand-rolling tobacco, but mainly hand-rolling tobacco	49%	1%
Hand-rolling tobacco only	0%	0%

F.19   The **upper bound** of total cigarette consumption is calculated firstly by estimating consumption levels from smokers who only smoked cigarettes. This is added together with a maximum consumption of cigarettes that could be smoked by dual smokers.

F.20   The **lower bound** of total cigarette consumption is calculated firstly by estimating consumption levels from smokers who only smoked cigarettes. This is added together with a minimum consumption of cigarettes that could be smoked by dual smokers.

F.21   Prior to 2012, the GLF was used as the base estimate for cigarette consumption. This was supplemented with OPN data on dual smokers where this was added/subtracted to obtain the upper and lower bounds.

Legitimate consumption

F.22   Estimates of legitimate consumption include:

- UK duty paid consumption
- cross-border and duty free shopping.

UK duty paid consumption

F.23   Estimates of UK duty paid consumption are taken directly from returns to HMRC (clearance data) on the volumes of cigarettes on which duty has been paid, along with the actual amounts of money. In general these can be used directly; however, there is a complication around changes to the duty and VAT rates normally pre-announced by the government at Budget. Not all cigarettes cleared in the months immediately before a duty or VAT rate increase will be consumed before the increase takes effect, meaning that the figures do not match the consumption estimates. To correct for this, a monthly average is taken to cover the period before and after the date of rate increase.

Cross-border shopping and duty free

F.24   Estimates of consumption of goods purchased as cross-border shopping are based on data from the International Passenger Survey (IPS). This provides estimates of the number of cigarettes an average adult traveller brings into the country, separately for air and sea passengers. The IPS figures are weighted by the ONS, scaling up the survey data to represent the total cross-border shopping entering the UK.

F.25   This estimate, however, does not cover sales made on-board ferries, so commercially provided data about deliveries of cigarettes to ferries is used to supplement the cross-border shopping estimate.

F.26   Duty Free cigarettes brought into the UK are also estimated from the IPS, using passengers coming back from outside the EU.

F.27   Legitimate consumption can be summarised as:

| Legitimate Consumption | = | UK Duty Paid Consumption | + | Cross-Border Shopping | + | Duty Free |

Conversion to monetary losses

F.28   Revenue losses associated with the illicit market are then estimated by combining the illicit market share information with price data and duty and VAT rate information.

F.29   All calculations to this point have been made on volumes of cigarettes. Volumes are converted to estimates of revenue using:
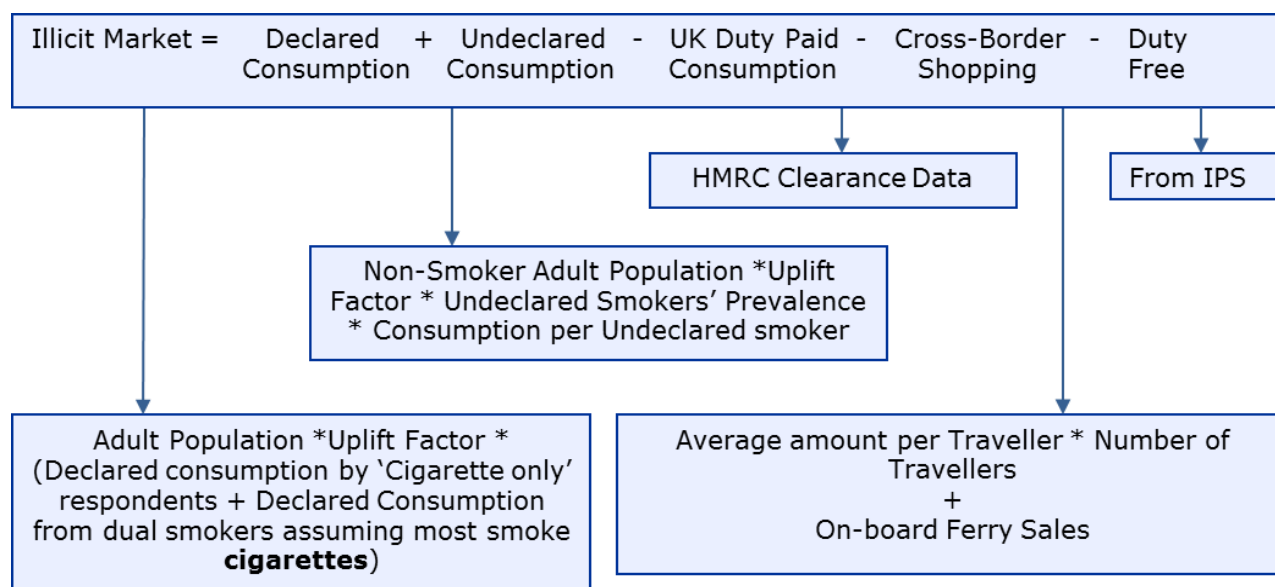
Losses = (Specific Duty + (Ad Valorem Duty + VAT Fraction) * Average Price) * Illicit Volume

F.30   The average price is taken as the weighted average price (WAP) of all cigarettes that were UK duty paid. The WAP is calculated by weighting the retail price of each product by the share of clearances in the cigarette market.

F.31   The VAT fraction is the proportion of the retail price that is VAT – for example, a 20% VAT rate is equivalent to a 1/6 VAT fraction. VAT fractions are calculated annually to capture changes in the VAT rate. This method assumes that VAT is also lost on all purchases. In some cases, the final illicit product is sold in legitimate outlets this may not always be the case, and this will be an overestimate of losses.

Summary of methodology

F.32   A summary of the calculation of the illicit market **Upper Bound** is:

| Illicit Market = | Declared Consumption | + | Undeclared Consumption | − | UK Duty Paid Consumption | − | Cross-Border Shopping | − | Duty Free |

HMRC Clearance Data

From IPS

Non-Smoker Adult Population *Uplift Factor * Undeclared Smokers' Prevalence * Consumption per Undeclared smoker

Adult Population *Uplift Factor * (Declared consumption by 'Cigarette only' respondents + Declared Consumption from dual smokers assuming most smoke **cigarettes**)

Average amount per Traveller * Number of Travellers
+
On-board Ferry Sales

F.33   A summary of the calculation of the illicit market **Lower Bound** is:

Illicit Market =	Declared Consumption	+	Undeclared Consumption	−	UK Duty Paid Consumption	−	Cross-Border Shopping	−	Duty Free

HMRC Clearance Data

From IPS

Non-Smoker Adult Population *Uplift Factor * Undeclared Smokers' Prevalence * Consumption per Undeclared Smoker

Adult Population * Uplift Factor * (Declared consumption by 'cigarette only' respondents + Declared Consumption from dual smokers assuming most smoke **hand-rolling tobacco**)

Average amount per Traveler * Number of Travelers
+
On-board Ferry Sales

Hand-rolling tobacco estimates

Methodology

F.34   The estimate of the illicit market for hand-rolling tobacco is produced using a top-down methodology as described in paragraphs F.1 to F.4.

F.35   Details of the estimation of total consumption and of legitimate consumption are provided in the subsequent sections.

Total consumption

F.36   The total consumption in any given year is calculated using the following:

- estimates of prevalence (proportion of the population that smokes hand-rolling tobacco) from the GLF, OPN and HSE

- estimates of hand-rolling tobacco consumption per smoker from GLF, OPN and HSE

- estimates of the adult population (16+) from the ON

- an uplift factor covering under-reporting.

F.37   The estimate of total UK consumption of hand-rolling tobacco for each year is based on the estimates of hand-rolling tobacco smoking prevalence and consumption per smoker for declared and undeclared smokers.

F.38   In general, most smokers admit that they smoke. The prevalence and consumption per smoker of these declared smokers was obtained from the GLF until 2011. As with the cigarette tax gap, from 2012 the OPN was used.

F.39   There are some smokers who, for whatever reason, do not admit that they smoke. The prevalence and consumption of Undeclared Smokers in the non-smoking population can be obtained from the HSE.

Uplift factor

F.40   The survey data for tobacco consumption is subject to under-reporting due to the self-reporting nature of the surveys. An uplift factor is necessary to correct for the bias. This uplift factor is calculated by taking estimates of consumption in a base year, and comparing these with independent estimates of total consumption. To do this we take a year in which there is believed

to be little or no illicit market and use HMRC clearance data, duty free and cross-border shopping estimates as a true indication of total consumption – for hand-rolling tobacco, 1984-85 is used.

F.41   Ideally, in order to reduce the impact of sampling error from the GLF, we would derive the uplift factor using the average of three years' data. However, the GLF in the 1980s only collected data on smoking prevalence in alternate years. This means that we do not have data for 1983-84 and 1985-86, and we have to rely on data from the single year 1984-85. This will increase random variation, and hence the resultant confidence intervals. However, figures for legitimate consumption do exist for the three years, so an average for legitimate consumption is calculated for 1983-84, 1984-85 and 1985-86. The uplift factor is defined as:

Uplift Factor =	Average Consumption in 1983-84 to 1985-86 (based on HMRC clearances and duty free/cross-border shopping)
	Total Consumption in 1984-85 (based on self-reported consumption from GLF survey)

F.42   Uncertainties in the data sets used to create these estimates mean that it is only possible to provide estimates in the form of a range within which the level of the illicit market is likely to fall. So, the exercise is undertaken twice – once to produce an upper bound for total consumption, and once to produce a lower bound. This allows us to produce a range for total consumption that takes account of the uncertainty in the underlying data.

Upper and lower bounds for total consumption

F.43   The only difference between the upper and lower bound consumption calculations is the treatment of dual smokers. In the upper bound calculation the majority of the dual smokers are considered to be cigarette smokers. In the lower bound estimate, we assume the majority of them smoke hand-rolling tobacco. This is explained further in the table below and the following section. Upper and lower bound assumptions:

OPN Survey Options	Allocation of total tobacco consumption to hand-rolling tobacco consumption	
	Upper bound assumption	Lower bound assumption
Cigarettes Only	0%	0%
Dual Smokers: Cigarettes and hand-rolling tobacco, but mainly cigarettes	49%	1%
Dual Smokers: Cigarettes and hand-rolling tobacco, but mainly hand-rolling tobacco	99%	51%
Hand-rolling tobacco only	100%	100%

F.44   The **upper bound** of total hand-rolling tobacco consumption is calculated firstly by estimating consumption levels from smokers who only smoked hand-rolling tobacco. This is added together with a maximum consumption of hand-rolling tobacco that could be smoked by dual smokers.

F.45   The **lower bound** of total hand-rolling tobacco consumption is calculated firstly by estimating consumption levels from smokers who only smoked hand-rolling tobacco. This is added together with a minimum consumption of hand-rolling tobacco that could be smoked by dual smokers.

F.46   Prior to 2012, the GLF was used as the base estimate for hand-rolling tobacco consumption. This was supplemented with OPN data on dual smokers where this was added/subtracted to obtain the upper and lower bounds.

Legitimate consumption

F.47   Estimates of legitimate consumption include:

- UK duty paid consumption

- cross-border and duty free shopping.

UK duty paid consumption

F.48   Estimates of UK duty paid consumption are taken directly from returns to HMRC (clearances data) on the volumes of hand-rolling tobacco on which duty has been paid, along with the actual amounts of money. In general these can be used directly; however, as with cigarettes, there is a complication around changes to the duty and VAT rates normally pre-announced by government at Budget. Not all hand-rolling tobacco cleared in the months immediately before a duty or VAT rate increase will be consumed before the increase takes effect, meaning that the figures do not match the consumption estimates. To correct for this, a monthly average is taken to cover the period before and after the date of rate increase.

Cross-border shopping and duty free

F.49   Estimates of the consumption of goods purchased as cross-border shopping are based on figures produced by the IPS. This provides estimates of the amount of hand-rolling tobacco an average adult traveller brings into the country, separately for air and sea passengers. The IPS figures are weighted by the ONS, scaling up the survey data to represent the total cross-border shopping entering the UK.

F.50   This estimate, however, does not cover sales made on board ferries, so commercially provided data about deliveries of hand-rolling tobacco to ferries is used to supplement the cross-border shopping estimate.

F.51   Cross-border shopping is only legitimate where it is for personal consumption. There is no clear, universal amount above which the goods are considered to be smuggled rather than for personal consumption, therefore we estimate upper and lower bounds.

F.52   The upper and lower bound estimates for cross-border shopping are calculated by changing:

- the definition of a smuggler
- whether or not to include non-respondents in the estimates.

Upper bound for cross-border shopping

F.53   The higher estimate for cross-border shopping has been created by assuming that:

- everyone who is coming back from EU Countries and agrees to answer the IPS questions is legitimately shopping and not smuggling
- all non-respondents will on average be carrying the same amount as those who complete the survey.

F.54   This leads to an estimate defined by:

Upper Cross-Border Shopping	=	High Average Hand-Rolling Tobacco per Traveller from IPS	*	Number of Travellers

Lower bound for cross-border shopping

F.55   The lower estimate for cross-border shopping has been created by assuming that:

- anyone who is carrying over 2.75 kg of hand-rolling tobacco is smuggling[2]
- all non-respondents are not carrying hand-rolling tobacco.

---

[2] Until 2011, the maximum amount that could be brought in from other EU member states was 3 kg. There is now no limit on this amount but this level has still been used for calculating the upper and lower bounds in the subsequent years.

F.56   This leads to an estimate defined by:

Lower Cross-Border Shopping	=	Low Average Hand-Rolling Tobacco per Traveller from IPS	*	Number of Travellers

F.57   Duty free hand-rolling tobacco brought into the UK is also estimated from the IPS, using passengers coming back from outside the EU.

F.58   Legitimate consumption can be summarised as:

Legitimate Consumption	=	UK Duty Paid Consumption	+	Cross-Border Shopping	+	Duty Free

Conversion to monetary losses

F.59   All calculations to this point have been made on volumes of hand-rolling tobacco. Volumes are converted to estimates of revenue losses using:
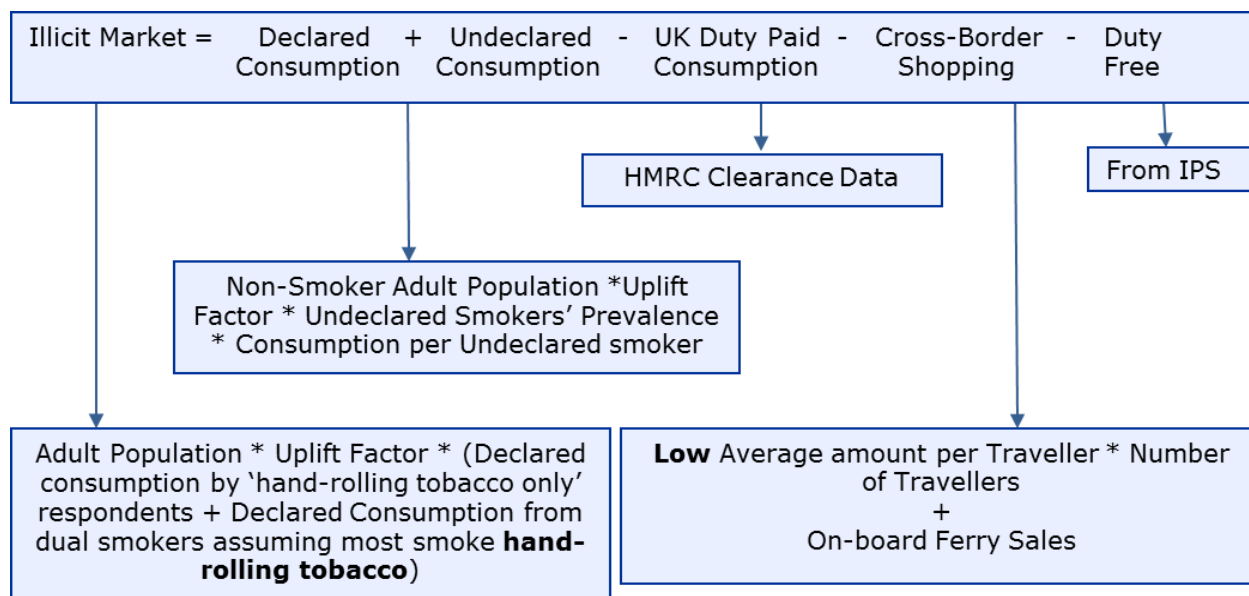
$$\text{Losses} = (\text{Specific Duty} + \text{VAT Fraction} * \text{Average Price}) * \text{Illicit Volume}$$

F.60   From 2012-13, the average price is taken as the WAP of all hand-rolling tobacco that was UK duty paid. The WAP is calculated by weighting the retail price of each product by the share of clearances in the hand-rolling tobacco market. WAP data for hand- rolling tobacco is not available prior to 2012-13 so the price of the most popular brand of hand-rolling tobacco was used.

F.61   The VAT fraction is the proportion of the retail price that is VAT – for example, a 20% VAT rate is equivalent to a 1/6 VAT fraction. VAT fractions are calculated annually to capture changes in the VAT rate. This method assumes that VAT is also lost on all purchases. In some cases, the final illicit product is sold in legitimate outlets this may not always be the case, and this will be an overestimate of losses.

Summary of methodology

F.62   A summary of the calculation of the illicit market **Upper Bound** is:

F.63   A summary of the calculation of the illicit market **Lower Bound** is:

| Illicit Market = | Declared Consumption | + | Undeclared Consumption | - | UK Duty Paid Consumption | - | Cross-Border Shopping | - | Duty Free |

HMRC Clearance Data

From IPS

Non-Smoker Adult Population *Uplift Factor * Undeclared Smokers' Prevalence * Consumption per Undeclared Smoker

Adult Population * Uplift Factor * (Declared consumption by 'hand-rolling tobacco only' respondents + Declared Consumption from dual smokers assuming most smoke **cigarettes**)

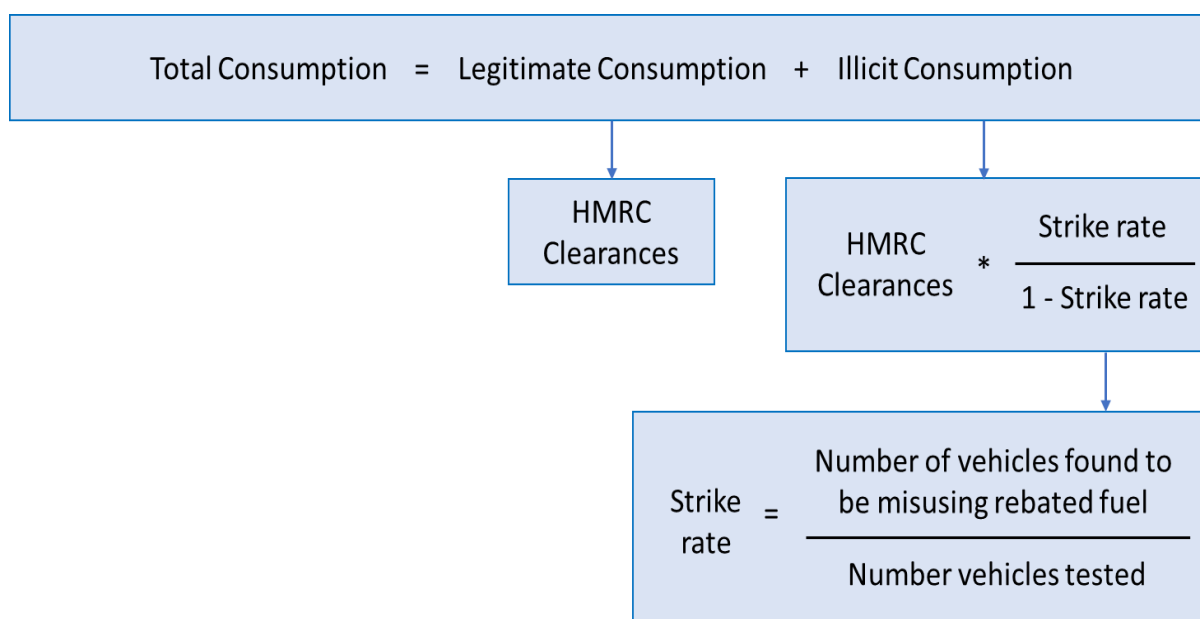**High** Average amount per Traveller * Number of Travellers
+
On-board Ferry Sales

Methodology

G.1 Previous estimates of the diesel tax gap were calculated using a top-down methodology. This methodology was no longer fit-for-purpose from 2013-14 as it was not sensitive enough to accurately measure the low tax gap. This meant that the estimates for 2013-14 were rolled forward for 2014-15 and 2015-16.

G.2 A new bottom-up methodology has been introduced to estimate the diesel tax gap for 2016-17. This means that the figures for 2016-17 are not directly comparable to previous years. The Great Britain (GB) and Northern Ireland (NI) tax gaps are calculated separately but the methodologies are identical.

G.3 Summary of methodology:

- legitimate consumption is based on the returns that HMRC receives from the volumes of diesel on which duties have been paid (HMRC clearances)

- illicit consumption is estimated using the proportion of vehicles found to be misusing rebated fuel in a random sample survey conducted by HMRC in 2017

- revenue losses (gross tax gap) associated with illicit consumption are estimated using average retail prices, duty rates and VAT rates

- the net tax gap is then calculated as the gross tax gap minus compliance yield.

Estimating total consumption

G.4 In April to June 2017, HMRC conducted a random survey where vehicles were stopped at the roadside and tested for illicit diesel. A stratified sample of 1,900 vehicles across the UK (1,500 in GB & 400 in NI) was used. The sample was stratified by vehicle type and region to ensure the results were representative.

G.5 The proportion of vehicles found to be misusing rebated fuel (strike rate) in this survey is used as an estimate of the proportion of vehicles misusing rebated fuel in the UK. The strike rate is then applied to legitimate consumption to give an estimate for illicit consumption. A separate strike rate is calculated for GB and NI.

G.6 The calculation of total diesel consumption can be summarised as:

$$\text{Total Consumption} = \text{Legitimate Consumption} + \text{Illicit Consumption}$$

$$\text{HMRC Clearances}$$

$$\text{HMRC Clearances} * \frac{\text{Strike rate}}{1 - \text{Strike rate}}$$

$$\text{Strike rate} = \frac{\text{Number of vehicles found to be misusing rebated fuel}}{\text{Number vehicles tested}}$$

Conversion to monetary losses

G.7   The diesel tax gap is driven by the misuse of rebated fuel. Rebated fuel is subject to a lower duty rate and has a lower retail price including VAT. Revenue loss occurs where this fuel is misused, and so should have been subject to a higher rate of fuel duty and additional VAT.

G.8   In order to estimate the revenue losses associated with the misuse of rebated fuel, the duty and VAT that has already paid needs to be taken into account. Therefore the difference between rebated and un-rebated duty rates has been used to estimate the duty loss associated with the illicit market.

G.9   Similarly, the difference in average retail prices for rebated fuel and un-rebated diesel has been used to estimate the VAT loss associated with the illicit market. Published data from the Department for Business, Energy and Industrial Strategy has been used to calculate average retail prices.

Confidence intervals

G.10   The upper and lower estimates correspond to confidence intervals that indicate the range where the true value of the illicit market may lie and arises due to random sampling error.

Exclusions

G.11   Smuggling of diesel is excluded on the basis that it is believed to be a minor issue, the scale of which isn't currently quantifiable. Cross-border shopping is excluded due to a reduced price difference between the Republic of Ireland and NI, meaning there is limited motivation for cross-border shopping activities. Revenue losses are assumed to be related to the misuse of gas oil (red diesel) only. The misuse of other fuels (for example, kerosene) have been excluded on the basis that this is believed to be a minor issue, the scale of which isn't currently quantifiable.

Chapter H: Estimates from random enquiry programmes, avoidance and hidden economy

H.1   This chapter provides detail on HMRC's random enquiry programmes and the methodology used to produce the avoidance and hidden economy tax gaps. This chapter covers all the approaches taken to produce income tax (IT), National Insurance Contributions (NICs) and Capital Gain Tax (CGT) tax gaps as well as the Corporation Tax (CT) gap from small business and mid-sized business (MSB) enterprises. The CT gap estimates for large businesses are described in Chapter I.

Random enquiry programme estimates

H.2   There are three direct tax random enquiry programmes which are used to produce tax gap estimates. They cover:

- Self Assessment (SA) individuals and small partnerships
- small business and MSB employers
- CT for small businesses and mid-sized businesses.

H.3   Random enquiry programmes allow HMRC to estimate the extent of under-declaration of liabilities arising from the submission of incorrect returns. Each return selected is subject to a full enquiry involving a complete examination of records. Under certain circumstances, a full enquiry may not take place if the return can be verified through third party information.

Populations and sampling

H.4   The sizes of the samples for the three programmes are shown in Table H.1 below.

Table H.1: Sample sizes for the Self Assessment, employer compliance and Corporation Tax random enquiry programmes

Self Assessment		Employer Compliance		Corporation Tax	
Tax return year	Sample size	Tax return year	Sample size	Accounting period ending in year	Sample size
2004-05	6,482	2004-05	1,649	2004-05	408
2005-06	5,834	2005-06	1,649	2005-06	419
2006-07	3,217	2006-07	1,649	2006-07	460
2007-08	3,219	2007-08	1,649	2007-08	492
2008-09	3,221	2008-09	1,649	2008-09	491
2009-10	2,599	2009-10	1,649	2009-10	480
2010-11	2,450	2010-11	825	2010-11	490
2011-12	2,599	2011-12	825	2011-12	567
2012-13	2,601	2012-13	925	2012-13	671
2013-14	2,451	2013-14	925	2013-14	540
2014-15	2,763	2014-15	925	2014-15	583
2015-16		2015-16	925	2015-16	362
2016-17		2016-17	925	2016-17	

H.5   To produce population estimates for total tax gaps from the samples in Table H.1, the average tax gap estimates from random enquiries are multiplied by the number of taxpayers in the population.

H.6 Adjustments are made to the population for cases deselected because they are outside of the population of interest– that is, no enquiry is made because criteria for not taking up an enquiry have been satisfied.

Self Assessment

H.7 The Self Assessment (SA) random enquiry programme allows us to estimate the tax gap arising from under-declaration of tax liabilities of individuals in SA. Results from the SA random enquiry programme are scaled up to the total number of individuals sent a SA notice to file.

H.8 In this context, 'individuals' means individuals who are self-employed, pensioners, and partnerships (with up to four partners), as well as those who are employees or may only have investment income. The taxes directly included are:

- IT
- NICs
- CGT.

H.9 The random sample used for the programme is selected from SA taxpayers issued with a notice to file a return. The sample is drawn by a systematic process that selects every "n^{th}" 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, we deselect that return. 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.

H.10 2009-10 is the last year which uses a simple random sample, as random samples for subsequent years have been stratified to improve the accuracy of the results. Samples drawn from SA business taxpayers are stratified by turnover from 2010-11 onwards, with samples drawn from SA non-business taxpayers stratified by level of income from 2011-12 onwards.

H.11 SA business consists of the self-employed and partnerships. SA non-business consists of employees, pensioners, trusts and all other types of SA taxpayers. In order to improve how representative the sample is, a weighting is applied based on how these customer groups are distributed across the population. We continue to review the customer group population assumptions.

H.12 Due to a relatively small sample size and large natural variance in the levels of under-declared liabilities from year to year, a smoothing approach has been used for small partnerships from 2010-11, when the stratification of business taxpayers was introduced. A three-year moving average with a double weighting given to the current year is used to smooth the data. This ensures that the resulting estimates are less susceptible to sampling variability and more indicative of longer-term trends.

Employer compliance

H.13 The employer compliance (EC) random enquiry programme allows us to estimate the tax gap arising from Pay As You Earn (PAYE) failures and other irregularities. Results from the EC random enquiry programme are scaled up to the total number of PAYE schemes.

H.14 The employer may be a self-employed individual, partnership, or a company and will be required to make returns under the PAYE regulations to account for IT and NICs.

H.15 The figures relate solely to IT and NICs collected through PAYE due on earnings and other income from employment. The scope of these figures also includes tax due on occupational pensions taxed through PAYE.

H.16 The taxes directly included are:

- IT
- NICs
- tax on occupational pensions.

H.17 The random sample is selected from using the former small and medium-sized enterprises (SME) customer classification and stratified on the basis of employer segments (defined in terms of the number of employees and whether the employer's business is incorporated). The resulting enquiries determine the small business tax gap and an element of the MSB tax gap.

Corporation Tax

H.18 The CT random enquiry programme allows us to estimate the tax gap arising from incorrect CT returns of small businesses. Results from the CT random enquiry programme are scaled up to the total number of live small business trader cases. In this context, 'live' excludes cases which are, for instance, dormant or dissolved/struck off.

H.19 For CT, the random sample is selected using the former SME customer classification from businesses which have been issued a notice to deliver a return. From April 2013, the sampling process changed to a stratified random sample, based on the size of annual trading turnover. This change allowed the CT random enquiry results to be weighted by the actual population of each strata resulting in an improved accuracy of the tax gap results. In addition to this, cases are excluded if the company has not submitted a return for the year of interest. The resulting enquiries determine the small business tax gap for CT and an element of the MSB CT gap.

H.20 For 'Measuring tax gaps 2018' the methodology regarding our treatment of deselected cases was updated. We now assign a proportion of deselected cases yield based on the average rate of under-declared liabilities for the strata the case falls in. This methodological change results in a large revision to the small business CT estimates from 2012-13 onwards.

H.21 Due to a relatively small sample size and large natural variance in the levels of under-declared liabilities from year to year a smoothing approach is used. A three-year moving average with a double weighting given to the current year is used to smooth the CT small business data throughout the series. This ensures that the resulting estimates are less susceptible to sampling variability and more indicative of longer-term trends.

Data Features

H.22 The latest observed random sample for SA used in the 'Measuring tax gaps 2018 edition' estimates are for 2014-15. More detail of the timing of random enquiries is given in the next section. An evaluation exploring the impact of changing the way of working SA business random enquiry cases to desk-based rather than face to face was began partway through the 2014-15 sample. This evaluation has continued into the 2015-16 tax year and is currently ongoing, with the results currently under review.

H.23 The latest observed EC random sample is for 2016-17. From 2015-16, approximately half of the sample was worked as a desk-based enquiry rather than the standard face to face approach. An evaluation of the effect of this change has been carried out and found no statistically significant evidence that it affected the outcome of the enquiry.

H.24 CT tax gap estimates use the latest observed random sample available which is 2015-16 for this year's publication.

Timing

H.25 There are two factors which influence the timing of the latest available tax gap estimate for a particular type of tax return:

- delays inherent in the returns process; this varies according to the head of duty and is shown in Table H.2 below

- delays due to the complexity of some random enquiries; it can take several years before sufficient random enquiries relating to a particular tax year are settled to robustly report the results.

Table H.2: Comparison of delays due to returns process

Random enquiry programme	Delays due to returns process
Self Assessment	Individuals generally have until 31 January following the year of assessment to which the return relates to submit their return. Once the return is submitted, HMRC then has a further year in which to open an enquiry.
Employer compliance	None. EC reviews initially look at the records of the previous 12 months.
Corporation Tax	Companies have until a year after the end of their accounting period to submit their return. HMRC then has a further year in which to open an enquiry.

H.26   There are three consequences of the timing issues described above:

- estimates of tax gaps for CT and SA are not available for the latest years due to a lag in data available. In order to present a more consistent picture of the scale of tax losses, projection factors have been applied to the estimates for CT and SA. We use the latest available data to project future years as this allows us to most effectively reflect recent policy and other changes that have a long term impact on taxpayer behaviour. These projection factors are shown below in Table H.3

- estimates for earlier years have been revised since previously published, as a result of the inclusion of additional data from reviews that have since been completed

- at the time of estimation, some enquiries were not closed for each year of each random enquiry programme. In order to estimate tax gaps 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 for such enquiries are made based on the results of recently settled enquiries with similar durations.

Table H.3: Comparison of projection factors

Random enquiry programme	Projection factors
Self Assessment	Based on the year-on-year changes in SA liabilities from 2014-15 onwards. Separate projection factors for SA business and SA non-business taxpayers are used.
Corporation Tax	Based on the year on year change in estimated total small business Corporation Tax liabilities from 2015-16.
Employer Compliance	No projection factor is used.

Sources of error

H.27   There are two main sources of error associated with the results of random enquiries which could result in the true values of the tax gaps differing from the estimates produced. These are:

- sampling variation in the 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; and

- systematic uncertainty where the sample results consistently tend to under-report the true values for the population, or where the sample does not include the full population, for example those participating in avoidance. We are able to make an adjustment for one source of systematic uncertainty, which is non-detection of non-compliance; details are discussed below.

H.28 The random enquiry programmes will not identify all incorrect returns or the full scale of under-declaration of liabilities, and so estimates produced from the unadjusted results of the programmes would underestimate the full extent of the tax gap. The Internal Revenue Service (IRS) in the United States has previously tackled this problem by using a range of 'multipliers' to adjust for non-detection[3]. The principles behind the IRS methodology have been applied to HMRC's data to produce approximate multipliers for the UK.

H.29 The IRS was able to undertake this analysis of non-detection because their random enquiry samples covered upward of 50,000 cases – much higher than is feasible in the UK. In the absence of this data for the UK, the US multipliers are used to account for non-detection. The size of the multipliers varies by the type of non-compliance found and are consistent year-on-year; Table H.4 shows how these multipliers differ by each random enquiry programme.

Table H.4: Comparison of adjustments for non-detection

Random enquiry programme	Adjustment to central estimate	Adjustment to lower estimate	Adjustment to upper estimate
Self Assessment (business)	1.908	1.000	3.075
Self Assessment (non-business)	1.260	1.000	1.928
Employer compliance	n/a	n/a	n/a
Corporation Tax	1.376	1.000	1.859

Modelling adjustments

H.30 From April 2013 the additional rate of income tax, charged on income above £150,000, was reduced from 50% to 45%. This policy change was thought to result in some taxpayers delaying income from 2012-13 to 2013-14 to take advantage of the lower rate (sometimes referred to as reverse forestalling).

H.31 In April 2016 the dividends taxation rates increased for all tax bands, and a £5,000 dividends allowance was introduced. This prompted many taxpayers (especially those in the Additional Rate band) to bring forward (or forestall) their dividend income to the 2015-16 tax year rather than paying the higher rates in 2016-17.

H.32 These changes in taxpayer behaviour have led to a decrease in tax liabilities in 2012-13 and 2016-17 and an increase in 2013-14 and 2015-16. The SA liabilities series used in this tax gap analysis have been adjusted to compensate for these changes in order to obtain a more accurate view of underlying liabilities in these years. The projection factors used to forecast liabilities to 2015-16 and 2016-17 have also been adjusted to allow for this effect.

Validation

H.33 As part of each year's programme, HMRC conducts a validation exercise for a sample of cases. These cases are checked to confirm that the enquiry outcomes (for example, the amount of yield) have been recorded accurately. Any inaccuracies are corrected prior to calculation of the tax gap for that year. Work has started to investigate whether the exercise can be modified to allow the

---

[3] James Andreoni, Brian Erard and Jonathan Feinstein (1998) 'Tax Compliance', Journal of Economic Literature, Vol. 36, No. 2. (June, 1998), pp. 818-860

correction of systematic errors to be projected onto the rest of the sample in a statistically valid way.

Outliers

H.34   Outliers are individual cases with large yields which are far removed from the yields of the other cases in the sample. Due to the nature of our samples our estimates are particularly sensitive to extreme values. To ensure that this small number of cases do not have an undue influence on the tax gap calculation their yield values are therefore capped. This allows us to use all valid information while smoothing the year-on-year variability.

H.35   Yield data is modelled using a representative statistical distribution. The final value used for each tax year is calculated as a 3 year moving average of the 99.85th percentile from this distribution, calculated based only on the results of years where the sample was stratified. For years before stratification, and years where a full 3 years of stratified results are not available, the value based on the last 3 complete stratified years is used.

H.36   A specific capping value is calculated for each random enquiry programme, including a separate value for SA business and non-business.

Deselections

H.37   Cases in the random enquiry programme are not worked for a number of reasons and this is done in a non-random way. This means that the cases which are not worked are likely to be systematically different from the cases that are worked. Cases which are not worked are called deselections or rejections. To avoid biasing the sample we treat and include cases that are deselected from the sample but are still within the population of interest. For these cases we impute a yield value based on the average yield and the proportion of under-declared liabilities of the strata it belongs to.

Tax gap calculation

H.38   The methodology used combines the estimate of under-declared liabilities with the amount of non-payment. As some of the tax gap is recovered through HMRC compliance activity, this is subtracted to give the net tax gap. The tax gap estimate is defined as:

> Net tax gap = (under-declared liabilities from incorrect returns * US multipliers to account for non-detection) + non-payment – yield from compliance activity.

H.39   The ranges which define the upper and lower estimates of the tax gap are based on the 95% confidence intervals of the estimate for under-declared liabilities from incorrect returns. These ranges are adjusted for non-detection as described in Table H.4 above.

Non-payment

H.40   The figures used to estimate levels of non-payment come from analysis of write-offs and remissions of tax on a financial year basis.

H.41   As separate figures of non-payment are not available for just the taxpayers within the scope of the random enquiry programmes, the amounts are split in proportion to the tax gap resulting from the relevant section of the populations. These non-payment figures will relate to the year when the loss was realised rather than the tax year the liability relates to. This approach has been taken because figures are not readily available by reference to the liability period.

Compliance yield

H.42   The random enquiries provide an estimate of the tax gap due to incorrect returns. However, HMRC carries out a wider programme of compliance activity to identify and correct erroneous returns. To calculate the net tax gap it is 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. See Chapter C.

Estimates for taxpayers not covered by the random enquiry programmes

Large and mid-sized employers operating a PAYE scheme

H.43   Larger employers with 250 or more employees, including those managed by our Mid-sized Business and Large Business directorates, are not covered by the EC random enquiry programme. This means an alternative methodology is required to produce an indicator of the associated tax gap.

H.44   An illustrative estimate is produced by assuming that the tax at risk will represent, over the long term, the same proportion of liabilities to small and the smallest of mid-sized employers, as shown by the results of the random programme. The estimated tax at risk is then adjusted to reflect compliance yield and non-payment.

H.45   The sustained fall in non-payments since 2014-15 is likely to have reduced the tax gap over those years. This fall has coincided with the introduction of the Real Time Information system, where information on payroll taxes is recorded more accurately and on a more frequent basis allowing us to identify debts and take action at an earlier stage than previously. The estimate of the tax gap has therefore been adjusted to incorporate the impact of this permanent change.

H.46   The estimated tax gap for small businesses usually represents between 1% and 2% of liabilities for this group on average (see Table 4.7 in 'Measuring tax gaps 2018 edition').

Large partnerships in Self Assessment

H.47   An illustrative estimate has been produced by assuming that the tax at risk will represent a similar proportion of liabilities to all other SA taxpayers, as shown by the results of the SA random enquiry programme. Projections for 2014-15 and 2015-16 are based on growth in large partnership liabilities.

Large and mid-sized businesses paying Corporation Tax

H.48   Tax gap estimates for large and MSB not covered by the CT random enquiry programme are covered in Chapter I.

Avoidance

Data sources

H.49   This section describes estimates of the avoidance tax gap for IT, NICs and CGT. The same data sources are used to estimate the avoidance tax gaps for CT small and mid-sized businesses, VAT and Stamp Duty Land Tax.

H.50   The avoidance tax gap is estimated using information that HMRC collects on tax avoidance schemes and records on its management information system. This includes avoidance schemes for individuals, trusts, partnerships and employers. The information that HMRC collects relates to disclosed and undisclosed schemes:

- for schemes disclosed under disclosure of tax avoidance schemes (DOTAS), information is captured during the following process: promoters of avoidance schemes that are covered by

the avoidance disclosure rules[4] must disclose any new schemes to HMRC when they are made available to potential users. Disclosures must contain sufficient detail for HMRC tax specialists to understand how the scheme works. For each disclosure, HMRC issues a scheme reference number to the promoter, and taxpayers who participate in the scheme are required to notify HMRC of the reference number on their tax return (described here as a 'notification')

- undisclosed schemes are identified by HMRC tax specialists through their compliance work.

H.51   When reviewing both disclosed and undisclosed avoidance schemes, tax specialists record an estimate of the 'tax under consideration' based on the relevant information relating to these ongoing enquiries. Any additional tax ('compliance yield') that is collected following completed enquiries is also recorded.

H.52   Detailed taxpayer-level data on avoidance schemes is available for large businesses and high net worth individuals. This enables comparison of the tax under consideration and compliance yield for an individual scheme user. On aggregate, data on completed enquiries provides a basis to estimate expected compliance yield from ongoing enquiries.

Methodology

H.53   The tax gap is calculated by subtracting estimated compliance yield from tax under consideration:

Estimated tax gap   =   tax under consideration – estimated compliance yield

H.54   The **tax under consideration** estimate relates to ongoing and completed enquiries. For completed enquiries, an estimate of tax under consideration is calculated from the compliance yield figures. This is calculated by applying the ratio of the compliance yield to tax under consideration from the taxpayer-level data to the actual compliance yield data.

H.55   The **compliance yield** that is likely to be recovered for those under investigation is estimated using the ratio of the compliance yield to tax under consideration. This ratio is derived from the taxpayer-level data on completed avoidance enquiries.

H.56   Improved data has been used to assign the total tax under consideration to tax years to produce an improved annualised tax gap estimate.

Data quality

H.57   The main source of error in these estimates is that HMRC may not identify all avoidance schemes – which will lead to an underestimation of the tax gap. It is difficult to quantify the extent to which this source of error impacts upon the estimates.

H.58   There are a number of issues with the methodology to estimate the avoidance tax gap. These include:

- estimates of tax under consideration are made by tax specialists using all the information available at the time. As this information improves over time, the view of tax under consideration may change

- the ratio of compliance yield to tax under consideration will change over time as more enquiries are completed. Any difference between estimated compliance yield from ongoing enquiries and actual compliance yield will lead to revisions in the estimates

- there is no tax year attached to the 'tax under consideration'. Therefore, the distribution of scheme uses across tax years is used to derive annualised estimate

---

[4] http://www.hmrc.gov.uk/aiu/summary-disclosure-rules.htm

- CT avoidance for LB groups are excluded from the calculations to avoid double-counting with the separate avoidance estimate for these businesses. Any re-classification of users following better information would lead to revisions of the CT avoidance estimate.

H.59   As a result of these factors, the figures presented in the document are likely to be revised as more information becomes available.

H.60   The data on avoidance schemes are reviewed by HMRC analysts for consistency and accuracy. Over time, as the scope, quality and quantity of the data improves, HMRC will seek to improve the avoidance tax gap estimates.

Hidden economy

Moonlighters

H.61   Moonlighters are defined as individuals who are employees in their legitimate occupation but do not declare earnings from other sources of income. There are two separate methodologies for different parts of the moonlighters estimate: one for earned income – that is individuals whose undeclared source of income is from employment – and one for unearned income – that is non-employment undeclared additional sources of income.

H.62   For calculating earned income, data from the Hidden Economy Quantitative Survey (HEQS) was used. The survey was commissioned by HMRC in 2015 to understand the nature of the hidden economy and the characteristics of those involved. Data on prevalence and income from hidden economy activities was captured as part of this research. In total, 9,640 respondents were surveyed.

H.63   The estimate for unpaid tax on moonlighters' earned income in the survey's sample is calculated by subtracting the tax paid on declared income from the tax that would have been due on their earnings if they had declared all their income. This covers income tax and National Insurance Contributions (NICs), with allowances made for whether the hidden economy activity in question would be classified as self-employment or employment. An allowance for under-reporting of income is also made in line with academic literature.

H.64   This sample estimate is then grossed up to the total population by using the prevalence rates of moonlighters with earned income in the population. These prevalence rates are obtained from the HEQS, and include weighting for non-response so that the prevalence rates are representative of the overall population.

H.65   A time series for the moonlighters' earned income estimate was created by using a proxy index which took into account changes in receipts over time as well as data from the Family Resources Survey (FRS). The FRS is a government sponsored study which provides information about households in the UK.

H.66   The tax gap for moonlighters' unearned income covers those individuals who have additional sources of income that are not from employment. These sources of income would therefore require them to submit a Self Assessment (SA) return to complement their normal tax payment through PAYE.

H.67   The sources of income covered by unearned income are lettings, interest, capital gains on property, chargeable events, Individuals Savings Accounts (ISAs) and secondary income (for example, activities such as hobbies or online selling that are not regular enough to be considered employment).

H.68   It is not necessary for most taxpayers to submit a SA 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. However, there are risks within this population, for example due to taxpayers not informing HMRC about sources of income, especially where they may exceed tax-free allowances. Where a SA return should have been completed, lettings, interest and ISA income would be subject to income tax; capital gains on property and chargeable events would be

subject to Capital Gains Tax (CGT); and secondary income would be subject to income tax and NICs.

H.69   HMRC cannot conduct random enquiries into the tax affairs of individuals who did not file a return because the legal position requires a return to be filed for an enquiry to take place. An alternative method is required for measurement of risks and estimating the associated tax gap.

H.70   HMRC has therefore used data matching of administrative data and third party information to measure the extent to which taxpayers fail to declare these additional sources of unearned income, with an estimate of additional tax due being calculated from the identified undeclared income. Third party data matched with administrative tax records includes rental deposit schemes and bank and building society interest declarations. Because of the large amount of data involved in this exercise, data matching is only conducted on a representative sample of the population already in PAYE. The results are thereafter grossed up from the sample to produce an estimate of the overall tax gap from moonlighters' unearned income.

H.71   The limitations associated with the results of this exercise relate to the coverage of the third party data used to establish evidence of additional undeclared income. Coverage varies across different sources of income, being especially good for lettings and interest income, whereas it is less reliable for the remaining sources identified. Additionally, there are other sources of income that could not be investigated due to unavailability of data. The resulting estimate should be interpreted broadly as a lower limit for the true scale of the tax gap relating to this group of taxpayers.

H.72   The latest estimate of the tax gap relating to moonlighters' unearned income is for 2014-15. This is projected forward based on receipts changes over time taking into account policy changes. For example, lettings income is subject to income tax; we take the lettings data-matching estimate for 2014-15 and multiply it by a value which adds together income tax receipts for 2014-15 and policy changes affecting receipts in 2015-16 to obtain an estimate of how much policy changes have increased or decreased income tax take by. This allows the projections to take into account changes in both tax rates and the tax base over time. For example, increases in the personal allowance reduce the potential tax revenue from hidden economy activities, all else being equal. The projections are based on the Office for Budget Responsibility's certified costings estimates for all income tax, NICs and CGT policy measures, and the relevant tax regime is applied for each of the unearned income sources.

H.73   Table H.5 shows the sources of income contributing to the estimated tax gap relating to moonlighters' unearned income for the 2014-15 data matching exercise.

Table H.5: Estimated tax gap for moonlighters' unearned income by source, 2014-15 (£ million)

Source of income	Estimated tax gap
Lettings	590
Interest	10
Capital gains	20
Irregular ISAs[1]	10
Chargeable events[2]	30
Secondary income	20
Total[3]	**680**

1 Irregular ISAs: Individuals are allowed to open or subscribe to just one ISA of any particular type within a given tax year. If another ISA of the same type is opened or subscribed to, then it is not entitled to tax-free status.
2 Chargeable events: chargeable events are most commonly payment of money from a life insurance policy, though other types of event such as the sale of an investment bond can also give rise to chargeable event gain that should then be treated as income for income tax purposes.
3 Figures rounded to the nearest £10 million. As a result components may not appear to sum.

Ghosts

H.74   Ghosts are defined as individuals who do not declare any of their income to HMRC, be it earned or unearned.

H.75   Data from the HEQS is used to estimate the ghosts tax gap. See the moonlighters section for details.

H.76   The estimate for unpaid tax on ghosts' income in the survey's sample is calculated by applying the relevant tax rate to the undeclared income estimated from the survey observations. This covers income tax and NICs, with allowances made for whether the hidden economy activity in question would be classified as self-employment or employment. An allowance for under-reporting of income is also made in line with academic literature.

H.77   This sample estimate is then grossed up to the total population by using the prevalence rates of ghosts in the population. These prevalence rates are obtained from the HEQS, and include weighting for non-response so that the prevalence rates are representative of the overall population.

H.78   As with moonlighters, a time series for the ghosts tax gap estimate was created by using a proxy index which took into account changes in receipts over time as well as data from the FRS.

Chapter I: Corporation Tax for large and mid-sized businesses

I.1   This chapter provides details on the methodology used to produce the Corporation Tax (CT) gap estimates for large businesses as reported in 'Measuring tax gaps'. The small and medium-sized enterprises (SME) CT tax gap estimates are described in Chapter H.

Businesses managed by Large Business

Overview

I.2   This section describes analysis of the CT tax gap for businesses managed by HMRC Large Business (LB). LB was formed in 2014 to manage tax compliance of the UK's 2,100 largest businesses.

I.3   Prior to April 2014, HMRC managed large businesses in two groups – the largest 800 groups in what was known as Large Business Service (LBS) and the next 10,000 most complex businesses were managed by Large and Complex (L&C). As such, estimates for large businesses combined both elements. In 2014, HMRC reorganised its structure, so that Large Business (LB) manage the largest 2,100 groups; the remaining businesses previously in L&C are now in mid-sized business (MSB).

I.4   For each LB business, HMRC tax specialists identify possible issues for further consideration with the business through a variety of approaches. The approach taken will depend on the relationship with the business, but where appropriate will include a detailed review of its accounts and CT return. The initial estimate of the amount of tax associated with these issues is recorded on the LB case management system as the tax at risk.

I.5   This initial estimate of tax at risk forms the basis for engagement with the business and further clarification of the issues. The initial estimate may be revised up or down when further information becomes available - these will lead to revisions to the tax gap estimates.

I.6   The tax gap is calculated as the difference between tax at risk and the compliance yield for avoidance risks and technical risks subject to litigation plus an uplift factor.

> Estimated tax gap = tax at risk – estimated compliance yield + uplift factor

Quality assurance

I.7   To complement the quality assurance practices in the LB, the accuracy of tax at risk was reviewed for a sample of large risks. The exercise showed that a small proportion of estimates required adjustment to reflect the latest known information. In addition, long-running cases tend to be more accurate.

Recording risk

I.8   The estimates are derived using information held on the LB case management system. Its primary purpose is to support effective management of risk and resource in the LB. However, because it allows managers and staff to record and monitor the tax at risk for businesses managed by the LB, the information it contains can be used to inform estimates of potential CT tax gap.

I.9   HMRC's case management systems are regularly reviewed to ensure that the most useful information is being collected. A significant change was made during 2011 and 2012 to transfer management information about compliance risks onto a new system. This resulted in tax at risk being captured earlier in the risk assessment process than before.

I.10   The tax at risk is an estimate of the maximum potential additional tax liability in each case before we have carried out a full investigation of the specific facts or analysis of relevant law. It is not actual tax either owed or unpaid, it is a tool to guide our enquiries to focus on the most significant risks that exist at any particular time with the largest businesses. In many cases, when we have

looked at the full facts it becomes clear that there is some lesser liability or even no further liability at all. Tax at risk will naturally vary from time to time as outstanding issues are settled and new risks are identified. The total is just a snapshot of work in progress and will naturally fluctuate as risks are settled and new ones taken up.

I.11   The tax specialists will engage with the business to gain a full understanding of the issue and to discuss the analysis of the tax position. The outcomes of the engagement with the business which might include formal review or alternative dispute resolution could be:

- HMRC agrees no additional tax is due

- agreement is reached between HMRC and the business that additional tax is due

- HMRC and the business are unable to reach agreement in which case the issue will be resolved by litigation.

I.12   The main cause of the net tax gap is where HMRC unsuccessfully challenges avoidance or loses in litigation cases. There will also be situations where issues are not identified and so HMRC does not clarify the situation with businesses. This second cause of the tax gap will not be captured on the case management system.

Classifying risks

I.13   For analysis of the tax gap, the recorded risks in HMRC's case management system are included in 'Measuring tax gaps' only if they are deemed to be non-compliant. Risks are categorised into avoidance and others.

I.14   The avoidance category relates to the use of disclosed avoidance schemes or other suspected avoidance identified by HMRC tax specialists. The avoidance disclosure regime was first introduced in 2004 and initially covered only specific aspects of CT. In August 2006 the regime was extended to the rest of CT. 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 the use of a scheme in their tax return. This provides an enhanced view of the use of marketed avoidance schemes and the potential associated tax at stake.

I.15   Other risks include a wide range of issues, from cases where there is genuine uncertainty about the correct tax treatment, through mistakes to culpable errors in, or omissions from, the company tax return. For risks identified that have fixed penalties, they are assumed to be settled for the right amount.

Open risks

I.16   It may take many years for enquiry work on issues to be completed and for the yield from this work to be recovered. So the likely compliance yield has to be partly forecast to derive figures for the tax gap. While all these issues will be closed over time, in the interim it is necessary to estimate the yield that will result from open enquiries.

I.17   The forecast of expected yield for such open cases is based on the results of closed cases from previous accounting periods. The amount of yield as a proportion of tax at risk that has resulted in yield in these cases is used to calculate the expected yield from the open cases. Once the actual results are known the estimates will be revised.

Risk working

I.18   The estimates for tax at risk are split into four main categories according to how the risks are worked by HMRC:

No net tax gap

a) Tax at risk for technical risks where agreement has been reached

b) Tax at risk for technical risks where agreement is expected

Estimated tax gap   = tax at risk – actual compliance yield (closed cases) – expect compliance yield (open cases)

c) Tax at risk for avoidance risks where the enquiry is closed or the risk is being worked

d) Tax at risk for technical risks that have been litigated or are leading to litigation.

Allocating risks to accounting periods

I.19   It is assumed that risks should relate to a maximum of five accounting periods. This is based on analysis of previous risks showing that around 90% of the risks relate to five or fewer accounting periods. This approach applies for the tax gap estimates from 2008-09 and onwards. As a result, the total number of risks from 2008-09 onwards are not directly comparable to the number of risks before 2008-09.

I.20   The estimates increase the understanding of the potential tax gap in businesses managed by the LB. As more data becomes available for analysis over time, improvements to the methodology can be made based on trends not apparent before. For example, since 2015 we have evidence from the LB Risk Task Force to derive an uplift factor for unidentified risks. The uplift is phased in to reflect the creation of LB Risk Task Force.

I.21   We have reviewed and improved the categorisation of international risks challengeable under the UK law. The inclusion of these risks has minimal impact on the tax gap because the compliance yield associated with these risks largely offsets the tax at risk.

I.22   The 2014-15 estimate is projected to 2015-16 and 2016-17 based on the trend in CT liabilities for groups formerly in the LB.

I.23   HMRC will be reviewing the methodology it uses to estimate the LB CT tax gap to ensure the best available data and assumptions are used. Therefore the estimates are subject to further revisions next year.

Mid-size businesses

I.24   Tax gaps for mid-sized business remain illustrative as it is difficult to estimate the tax gap as a whole from existing data sources and they are estimated using an experimental methodology. Random enquiries data can only be used to draw inferences on the level of non-compliance for the smallest businesses in this segment, and the remaining businesses are unlike LB, in which close scrutiny of every business means that most risk are identified.

I.25   The 2014-15 estimate is projected to 2015-16 and 2016-17 based on the trend in MSB CT liabilities. This produces estimates of tax at risk for 2015-16 and 2016-17, from which compliance yield is subtracted and an estimate of losses from non-payment is added.

Chapter J: Other Taxes

J.1   Other taxes include:

- Other direct taxes
    - Stamp Duty Land Tax
    - Stamp Duty Reserve Tax
    - Inheritance Tax
    - Petroleum Revenue Tax[5].

- Other indirect taxes
    - Customs duties and levies
    - Air Passenger Duty
    - Insurance Premium Tax
    - Climate Change Levy
    - Aggregates Levy
    - Landfill Tax.

J.2   With the exception of Stamp Duty Land Tax and Landfill Tax, tax gap estimates for all other taxes are illustrative and produced using management assumptions.

Stamp Duty Land Tax

Methodology

J.3   The Stamp Duty Land Tax (SDLT) tax gap is estimated using a combination of management information and management assumptions.

Tax under consideration

J.4   The SDLT tax gap is calculated from the amount of SDLT outstanding, referred to here as tax under consideration or tax at risk (TAR). There are four identified components to the amount of SDLT tax at risk (why the SDLT tax gap exists):

- TAR from cases being investigated by the SDLT Anti-Avoidance team

- SDLT avoidance unknown to the department

- reliefs that are improperly claimed

- SDLT not paid due to evasion, goodwill, agent behaviour and linked transactions.

SDLT avoidance unknown to the department

J.5   It would be impossible for HMRC to know about every case of SDLT avoidance, because either the associated paperwork has not been completed, or because it has been deliberately falsified and not yet discovered, or for some other reason. Expert opinion has suggested that HMRC is likely to be aware of approximately 80% of all transactions involving SDLT where tax under consideration has resulted. For this reason, a multiplier of 1.25 (100 / 80) has been used to 'uplift' the amount of tax under consideration which is known about to account for this.

Evasion

J.6   This reflects a percentage of the total amount of SDLT receipts (as published by HMRC) not initially paid because of evasion. Internal discussions with subject matter experts suggested that this amounts to 1% of the published SDLT receipts each year, with around 50% of this recoverable in line with other non-avoidance activity.

---

[5] Note that Petroleum Revenue Tax was zero-rated from 1 January 2018.

Reliefs improperly claimed

J.7   Improperly claimed relief takes different forms and there are more than 30 different reliefs claimed for SDLT. Four of these reliefs alone accounted for over 75% of the value of reliefs claimed and are used as the basis of this calculation.

J.8   Analysis of open enquiries and a series of pilot research projects have suggested that up to 5% of these claims may be falsely claimed. Additionally, there is an assumption that HMRC may only be able to recover 10% of the tax under consideration involved in these cases: this takes into account the large number of reliefs for which compliance work has not yet begun and the small number of cases open into those reliefs that have been targeted.

Goodwill, agent behaviour and linked transactions

J.9   This reflects a percentage of the total amount of SDLT receipts (as published by HMRC) not initially paid because of goodwill, agent behaviour and linked transactions. Internal discussions with subject matter experts suggested that this amounts to 0.5% of the published SDLT receipts each year, with around 50% of this recoverable in line with other non-avoidance activity.

Exclusions from this methodology

J.10   Estimates for years prior to 2011-12 include the amount of SDLT avoided by the use of tax avoidance schemes. These were artificial structures solely constructed to avoid SDLT that the department was aware of. This was calculated by multiplying together the number of disclosure of tax avoidance schemes (DOTAS) schemes, the estimated tax under consideration each year and the estimated number of users of each DOTAS scheme. This is excluded from 2011-12 onwards as it is believed that no further DOTAS schemes have been revealed to the department.

J.11   Estimates for years prior to 2015-16 include threshold manipulation (another form of SDLT evasion). This occurred when a sale value of a property was artificially reduced to below a threshold in order to reduce the SDLT liability. Previously, SDLT was charged at a single rate based on the value of the total purchase price. From 4 December 2014, SDLT liabilities changed to incremental rates applied only to the portion of the purchase price that falls within each rate band. This significantly reduced the potential value of tax lost due to threshold manipulation. For this reason, estimates after this point do not include threshold manipulation.

Landfill Tax

Methodology

J.12   The Landfill Tax gap is estimated using a combination of modelling, proxy indicators and assumptions made in collaboration with HMRC's operational experts. It uses HMRC and publically available data to estimate each component. From 1 April 2015, Landfill Tax was devolved to Scotland hence, since 'Measuring tax gaps 2017 edition', Scottish Landfill Tax is no longer in scope of this estimate. Landfill Tax attributable to Scotland is removed from the tax gap estimate by using the percentage of total UK Landfill Tax receipts attributable to Scotland.

Tax in scope

J.13   Landfill Tax is due on waste disposed of at a permitted landfill site as a disincentive to landfilling and to encourage better waste management. The tax gap measures the difference between the amount of Landfill Tax that should theoretically be collected, when waste is disposed of at a landfill site, and the amount that is actually collected.

J.14   The methodology does not include waste disposed of at illegal waste sites, as this is not taxable for the years included in this publication. This became taxable from 1 April 2018 so will be included in future publications for years 2018-19 onwards. There is also additional tax at risk from illegally disposed waste depending on what proportion, if any, would have been disposed of at landfill if it had not been illegally dumped. This has been excluded due to the extent of assumptions that would be required to estimate the tax at risk on illegal waste, and the inherent uncertainties around whether the waste would be disposed of by way of landfill.

Tax under consideration - under-declaration

J.15   Under-declared waste is estimated in two ways and averaged to arrive at a central estimate.

J.16   In the first method a trend line is fitted to HMRC data on taxable tonnes over time, then expected and actual tonnages of waste are compared. The estimate is refined to take account of the increase in diversion of waste away from landfill in recent years to incineration and export of refuse derived fuel. We assume nearly all of this diverted waste is taxable at the standard rate if sent to landfill.

J.17   After these adjustments, the tax under consideration is estimated by applying the tax rates at the same composition as declared taxable waste. The ratio of standard rate to lower rate has changed over time with it becoming roughly 50:50 in recent years.

J.18   In the second method, a proxy indicator is used to estimate under-declaration. This assumes that all landfill site operators have under-declared taxable waste by 5% per year, and that this under-declared amount should be taxed at the standard rate.

Tax under consideration - misclassification

J.19   There are two rates of Landfill Tax, standard and lower rate. A trend line is fitted to HMRC published statistics on lower rated tonnes declared over time. Expected tonnages of lower rate waste is then compared with declared lower rate waste. Declared lower rate waste shows a trend towards increasingly larger amounts of lower rated waste going to landfill in recent years. Some of this is expected due to changes in how waste is diverted away from landfill towards other forms of waste management.

J.20   We assume 25% of the difference between expected and declared lower rated waste constitutes the tax base under consideration. The tax under consideration is then the difference between the standard and lower rates of waste on this tonnage.

Tax gap calculation

J.21   As some of the gross tax gap is recovered through HMRC compliance activity, this is subtracted to give the net tax gap. The net tax gap estimate is defined as:

> Net tax gap = under-declared waste + misclassified waste (as either standard or lower-rated) – yield from compliance activity

Chapter K: Customer groups

Introduction

K.1   For the first time, the figures presented in 'Measuring tax gaps 2018 edition' adopts the customer groups applied by HMRC from April 2014.

K.2   HMRC segments its customers by type and size, and tailors its customer services based on behaviours, capabilities and the level of risk. This enables HMRC to use systems, processes and services that are closely tailored to the requirements of each group, while maintaining fairness and consistency across the board and addressing cross-cutting behaviours. Criminals are dealt with outside of these groups and are subject to a very different approach.

K.3   For the period up to 2014-15 HMRC's customer group breakdown was:

- large businesses, managed by HMRC's:
    - a) Large Business Service (LBS), the 800 largest and most complex businesses operating in the UK
    - b) Large and Complex (L&C), the next 10,000 largest businesses operating in the UK
- small and medium enterprises (SMEs), the remaining businesses operating in the UK
- individuals.

K.4   In April 2014, HMRC reoriented its activity around five clear customer groups:
- large businesses
- mid-size businesses
- small businesses
- individuals
- wealthy.

Tax gap definitions

K.5   Large business: the 2,100 largest and most complex businesses with an annual turnover typically exceeding £200 million. Most large businesses are complex groups of companies.

K.6   Mid-sized business: around 170,000 businesses with turnover of between £10 million and £200 million or with 20 or more employees and around 30,000 public bodies and charities. They make up less than 4% of UK business but employ more than 19% of the UK's workforce.

K.7   Small business: more than five million businesses with turnover below £10 million and fewer than 20 employees. It also includes micro-businesses that have turnover below £2 million and fewer than 10 employees.

Customer group segmentation methodology

K.8   There have been no changes to the methodology for estimating each of the individual components of the tax gap. Instead, data has been used to apportion the different components of the tax gap to the new customer groups. This has allowed for a consistent approach over time, meaning that a time series is available from 2005-06 onward.

K.9   All of the businesses that were previously managed by the LBS are now part of the Large Business (LB) directorate. In addition, the largest 1,300 businesses previously in the L&C unit are now part of LB.

K.10   The remaining 8,700 business previously managed by the L&C were reorganised into the Mid-Sized Business strand of HMRC's Wealthy and Mid-Sized Business Compliance directorate, as were SMEs with a turnover above £10 million or more than 20 employees.

K.11   The remaining businesses are now referred as small businesses, and are managed by HMRC's Individuals and Small Business Compliance directorate.

K.12   Table L1 below shows the relationship between the HMRC customer groups and the breakdowns used in previous publications.

Table L1: Relationship between customer breakdowns used in previous publications and the current HMRC customer groups

Customer breakdown in previous publications	Customer breakdown previously used by HMRC	HMRC customer groups (2014-present)
Large businesses	Large Business Service	Large businesses
	Large and Complex unit	
		Mid-sized businesses
Small and medium enterprises	Small and medium enterprises	Small businesses
Individuals	Individuals	Individuals
		Wealthy

K.13   The conversion of the tax gap estimates between the breakdowns used in previous publications and the HMRC customer groups was done on the basis of historical data. Data on tax receipts by customer group over a number of years was used to estimate the proportion of the tax gap going from each of the previously used breakdowns to the HMRC customer groups.

K.14   No separate estimate currently exists for the wealthy customer group; it is covered in individuals.

K.15   Table L2 below compares the 2016-17 tax gap estimates by customer breakdowns used in previous publications and by current HMRC customer groups.

Table L2: 2016-17 tax gap estimates by customer breakdowns used in previous publications and by current HMRC customer groups

Customer breakdown previously used	Tax gap estimate (2016-17, £bn) by customer breakdown	Tax gap estimate (2016-17, £bn) by HMRC customer group	HMRC customer groups (2014-present)
Large businesses	9.5	7.0	Large businesses
		3.9	Mid-sized businesses
Small and medium enterprises	15.1	13.7	Small businesses
Individuals	3.4	3.4	Individuals
Criminals	5.4	5.4	Criminals
Total 2016-17 tax gap	**33**	**33**	**Total 2016-17 tax gap**

Chapter L: Abbreviations

BEIS: Department for Business, Energy and Industrial Strategy

BBPA: British Beer and Pub Association

CGT: Capital Gains Tax

CHIEF: Customs Handling of Import and Export Freight

CT: Corporation Tax

DOTAS: Disclosure of Tax Avoidance Schemes

EC: Employer Compliance

EFS: Expenditure and Food Survey

EU: European Union

EX46:   Beer Duty Return

FRS: Family Resource Survey

FSS: Family Spending Survey

GB: Great Britain

GDP: Gross Domestic Product

GHS: General Household Survey

GLF: General Lifestyle Survey

HEQS: Hidden Economy Quantitative Survey

HMRC: Her Majesty's Revenue and Customs

HSE: Health Survey for England

IPS: International Passenger Survey

IRS: Internal Revenue Service (United States)

IT: Income tax

L&C: Large and Complex

LB: Large Business

LBS: Large Business Service

LCF: Living Costs and Food Survey

MSB: Mid-sized business

NAAFI: Navy, Army and Air Force Institutes

NHS: National Health Service

NI: Northern Ireland

NICs: National Insurance Contributions

ONS: Office for National Statistics

OPN: Opinions and Lifestyle Survey

PAYE: Pay As You Earn

SA: Self Assessment

SDLT:	Stamp Duty Land Tax
SME:	Small and medium-sized enterprise
TAR:	Tax at risk
UK:	United Kingdom
VAT:	Value Added Tax
VTTL:	VAT total theoretical liability
W1:	Excise Warehouse Return
WAP:	Weighted average price

HM Revenue & Customs