

Vehicle Excise Duty Evasion Statistics: Notes and Guidance



Department
for Transport

Last updated: December 2013

1. Introduction

This document provides guidance on the [Vehicle Excise Duty Evasion Statistics](#). Comparable statistics in this series are available for 2007 onwards. They are published every 2 years by the Department for Transport.

These Notes and Definitions include:

1. Introduction
2. Survey design and methodology
3. Changes to the general method
4. Derivation of evasion estimates
5. Effect and treatment of misread registration marks
6. Strengths and weaknesses of the data
7. Uses of the roadside survey data

2. Survey design and methodology

The following method has been used for VED surveys from 2007 to 2013, with only relatively minor changes from year to year. These changes are summarised in section 2 below.

During June in each survey year, over 1 million valid sightings of vehicle registration marks were collected at sites around the UK by contractors working on behalf of the Department for Transport. Collection was conducted at 256 sites, distributed across the regions of Great Britain and within Northern Ireland.

In Great Britain, four sites (one of each of four different road types) were selected in each of the 49 police force areas outside London. In London, sixteen sites (four of each road type) were selected. The road types covered in each area were: built-up A roads; non built-up A roads; built-up minor roads; and non built-up minor roads. In addition, two motorway sites were chosen in each country or region, with the exception of London, where four motorway sites were selected. As well as these, 20 sites were selected in Northern Ireland to cover the different road classes and to give a large enough survey sample to give reliable results. These sites have remained the same since 2007 to make year-on-year results as comparable as possible.

At each site, the contractors were required to collect data for one twelve hour period on a weekday and for six hours on a weekend day. On weekdays, surveying was between 8.00 am and 2.00 pm and between 3.00 pm and 9.00 pm, while on weekends surveying was for one of these periods only. Weekend surveying was predetermined to give equal numbers of morning and afternoon counts and equal numbers of Saturday and Sunday counts; otherwise surveying was down to the discretion of the contractors. The precise location of each site was agreed between the survey

contractors and the Department for Transport.

Collection was carried out by positioning a video camera at the roadside and then using Automatic Number Plate Recognition (ANPR) software at the contractor's office to collect details of the observed vehicles' registration marks from the video footage. The ANPR software automatically records each vehicles' registration mark from the video footage. Each automatically produced registration mark was then manually checked against the video footage of that vehicle. The number plates for motorcycles were collected by enumerators at the roadside, as motorcycles do not have forward facing number plates.

The collected data were then returned to the Department where those registration marks in an invalid or foreign format or where the vehicle was noted as displaying trade plates were removed. The remaining records were passed to the Driver and Vehicle Licensing Agency (DVLA, GB only) and the Driver and Vehicle Agency (DVA, Northern Ireland) in order to identify which registration marks were licensed when they were seen and which were unlicensed. Once these matched data were returned, a further quality assurance check was made by comparing the ANPR-produced registration mark with the video image of all vehicles reported to be unlicensed. Any misread registration marks identified through this process were removed from the survey dataset (see also section 5 below).

The number of records included in the final survey dataset is given in the table below.

Survey size, 2007 to 2013

Tax Class	2007	2008	2009	2010	2011	2013
<u>Great Britain</u>						
Private and Light Goods	1,482,422	1,292,779	1,301,819	1,424,814	949,148	908,076
Goods	92,967	81,115	58,755	64,289	45,003	44,542
Motorcycles	9,166	7,124	7,921	9,582	6,706	6,937
Bus	16,601	14,806	15,618	17,091	10,818	11,005
Exempt	50,260	45,756	51,360	57,673	39,220	36,304
Other	2,456	1,061	1,584	3,872	2,052	1,939
All tax classes	1,653,872	1,445,997	1,437,057	1,577,321	1,052,947	1,008,803
<u>Northern Ireland</u>						
Private and Light Goods	56,058	71,710	77,796	83,037	56,683	83,829
Goods	3,121	4,345	3,263	3,126	2,396	2,948
Other	5,411	6,903	7,467	8,261	5,877	8,306
All tax classes	64,590	82,958	88,526	94,424	64,956	95,083

3. Changes to the general method

Sample size reduction from 2011 onwards

In line with Government guidelines to reduce costs from 2010/11 onwards, the Department worked with the survey contractor find ways of delivering a reduced cost survey whilst maintaining a high standard of data. In order to achieve this, all non-motorway sites in 2011 were surveyed in one direction only. Motorways continued to be surveyed in both directions. This resulted in the sample size falling from around 1.4-1.6 million to about 1.1 million vehicles. As all sites were still surveyed and there is no reason to expect a higher rate of evasion in one direction than another, the figures should be as reliable as in previous years. Similar sampling was maintained in 2013.

Method of recording motorcycle registration marks from 2013

Prior to 2013, the manual collection of motorcycle numberplates was done by the site technician manually noting down the number, or recording it on a voice recorder and transcribing it later. In 2013, digital cameras were used to record motorcycles at 16 sites. Because these sites were selected on the basis of having larger numbers of motorcycles, they accounted for almost half the motorcycle sample. The original manual methods continued to be employed on the other sites. The digital camera method was considered successful on the basis that (i) clear images resulted (ii) the images enabled full QA checking of apparent unlicensed cases (iii) fewer records were rejected as being 'obviously invalid' than in previous years (i.e. a motorcycle plate as recorded in the survey corresponding to some other vehicle type in DVLA records).

Revisions to the Northern Ireland evasion rates

An improved weighting methodology was introduced for the overall evasion rate in Northern Ireland in 2011, and retrospectively applied to all estimates from 2007, bringing the overall rates for Northern Ireland closer to those for Great Britain. The rates provided for the private and light goods vehicles and goods vehicles tax classes in Northern Ireland were not affected.

4. Derivation of evasion estimates

The exact methodology used to derive the estimates within this report is complex and is outlined in detail within a methodological review carried out by Southampton University. This review is available at: <https://www.gov.uk/government/publications/vehicles-statistics-guidance>. However, the key stages within each calculation are summarised below.

Derivation of rate of unlicensed vehicles in traffic

1. The rate of unlicensed vehicles in traffic for each tax class as observed at each survey site is calculated.

2. The rates in (1) are weighted by traffic flow information specific to that site, drawn from national traffic census sources.
3. These weights are used to calculate an average unlicensed rate for each tax class within each road type and region.
4. The rates in (3) are further weighted by traffic flow information specific to each road type and region.
5. These secondary weights are used to calculate an estimated unlicensed rate in traffic for each tax class.

Derivation of rate and number of unlicensed vehicles in active stock

1. The number of times each individual vehicle was seen within the survey is calculated.
2. These data are used within a statistical model based on the negative binomial distribution, using the frequency of repeat sightings, to estimate the relative mileage of licensed and unlicensed vehicles within each tax class. A table showing the results produced by this model is given in the table below. From 2010, the distribution used to estimate relative mileage for the *Goods*, *Motorcycles*, *Bus* and *Other* tax classes include the sightings from all ANPR-based surveys (i.e. 2007 and later). This provides a larger sample which produces a more robust relative mileage (and hence 'in-stock') estimate.

Relative mileage of unlicensed vehicles in Great Britain, 2013

Tax Class	Relative Mileage¹	Lower Confidence Limit	Upper Confidence Limit
Private and Light Goods	1.12	1.03	1.21
Goods	0.79	0.72	0.87
Motorcycles	0.75	0.56	0.96
Bus	0.57	0.45	0.71
Exempt	1.30	0.85	1.84
Other	0.63	0.39	0.93

1 Average mileage per unlicensed vehicle for every mile travelled by a licensed vehicle.

3. These relative mileage estimates are combined with the rates of unlicensed vehicles in traffic to produce an estimated rate of unlicensed vehicles in active stock for each tax class.
4. The rates calculated in (3) are used together with the known number of licensed vehicles to produce the estimated number of unlicensed vehicles in active stock within each tax class.

-
5. The numbers calculated in (4) are summed to give the overall estimated number of unlicensed vehicles in active stock.

Derivation of revenue lost from unlicensed vehicles

1. The average cost of a yearly licence within each tax class is calculated. These costs are given in the table below:

Average licence values by tax class, Great Britain, 2007/8 to 2011/12

Tax Class	2007/08	2008/09	2009/10	2010/11	2011/12	2013/14
Private and Light Goods	£163	£168	£167	£165	£181	£180
Goods	£608	£612	£579	£574	£586	£617
Motorcycles	£45	£46	£46	£47	£52	£55
Bus	£273	£271	£264	£262	£264	£266
Exempt	£0	£0	£0	£0	£0	£0
Other	£215	£205	£221	£236	£232	£240

2. The values from (1) are multiplied by the estimated numbers of unlicensed vehicles within each tax class to give the estimated revenue lost from unlicensed vehicles within each tax class.
3. The values calculated in (2) are summed to give the overall estimated revenue lost from unlicensed vehicles

Confidence Intervals

Where possible, 95 per cent confidence intervals have been provided for the most recent estimates. These are calculated using exact binomial confidence intervals as outlined in 'Tian et al, 2009, *A comparative study of confidence intervals for negative binomial proportion*. Journal of Statistical Computation and Simulation, 79(3), pp 241-249'. As they are exact confidence intervals the figures are asymmetrical. This means that the upper and lower limits are not equally spaced from the central estimate.

The confidence intervals give a range of values around the central estimate. These mean that if the survey was carried out many times (collecting a different set of number plates, perhaps at different locations or on different days), 95 per cent of the surveys would give an evasion rate between the upper and lower value. In practice, this can be interpreted to mean that it is highly likely that the actual evasion rate in the whole population (as opposed to estimated rate for just our sample) will fall somewhere between the upper and lower limits found in the published tables. This is by no means certain, though, and there is a 5 per cent chance that actual evasion rate does not fall within the confidence limits produced from this survey.

The only way to know the actual evasion rate for the whole population would be to record every

single vehicle movement on the roads throughout the year. Clearly this is not practically possible, either from a technical or financial point of view. Therefore the survey sample figures presented in the table are the best estimates that can be produced given the practical restrictions.

As the 'in-stock' calculation combines the 'in-traffic' figures and relative mileage figures, it is not possible to provide 95 per cent confidence intervals on the same basis. Therefore upper and lower limits have been estimated by multiplying the upper limit from the 'in-traffic' figures by the upper limit from the relative mileage figures, and the same for the lower limits for both estimates. These have been used in tables VED0201, VED0202 and VED0301.

5. Effect and treatment of misread registration marks

Effect of misreads on evasion estimates

The use of ANPR (Automatic Number Plate Recognition) technology for the vast majority of the survey provides a good understanding of the effect of misread registration marks on the survey results.

To establish the nature of this effect, random samples of 1-2,000 vehicle registration marks collected through the various surveys have been taken and, for each instance, the registration mark recorded manually compared to a still video image of the vehicle as provided through the ANPR system. Where a misread was found to have occurred, the corrected registration mark was compared to a copy of the DVLA database to establish whether the actual vehicle sighted was licensed or unlicensed when it was seen.

The results of these checks concluded that:

- a) The overall rate of misreads for ANPR-based sightings within the 2013 survey could be estimated at between 0.72% and 1.67%, based on a random check of 2,000 cases.
- b) The presence of misread registration marks, if left untreated within the survey data, would serve to artificially inflate the survey's evasion estimates. This is clearly demonstrated by the table below which shows the evasion rates that would be expected under different scenarios if misreads were left untreated within the survey data.

Expected evasion rate without treatment of misreads

Actual evasion rate	Misread rate				
	1%	2%	5%	10%	20%
0.5%	0.7%	0.9%	1.5%	2.6%	5.0%
1.0%	1.2%	1.4%	2.0%	3.1%	5.5%
1.5%	1.7%	1.9%	2.5%	3.6%	5.9%
2.0%	2.2%	2.4%	3.0%	4.0%	6.4%

Treatment of misread registration marks in ANPR-based surveys

As every registration mark collected through the ANPR software is manually checked by the contractor within their data validation processes, it is likely that the majority of misreads are corrected before the survey data are submitted to the Department. One of the quality standards required by the Department is that a random sample of the data supplied by the contractor are >98 per cent accurate (tested by the Department by visual inspection against the original image, supported by comparison with DVLA records). In practice a standard of 98.8 per cent was achieved for the 2013 survey.

However, the analysis shown above demonstrates that even these relatively small levels of misreads can significantly inflate evasion estimates and, therefore the Department carries out an additional stage of quality checking by comparing the registration marks and video images of all vehicles reported to be unlicensed after matching to the DVLA and DVA databases. Any registration marks found to have been misread through this process are removed from the survey dataset.

While this additional process removed those remaining misreads incorrectly matching the registration mark of a separate 'unlicensed' vehicle, it will not have removed those erroneously matching 'licensed' vehicles - creating the potential for under-estimation of evasion. However, the extent of this under-estimation is likely to be negligible, as shown by the table showing expected evasion rates after the treatment of misreads below:

Expected evasion rate after treatment of misreads

Actual evasion rate	Misread rate				
	1%	2%	5%	10%	20%
0.5%	0.5%	0.5%	0.5%	0.5%	0.5%
1.0%	1.0%	1.0%	1.0%	1.0%	1.0%
1.5%	1.5%	1.5%	1.5%	1.5%	1.4%
2.0%	2.0%	2.0%	2.0%	2.0%	1.9%

In the case of motorcycles, data were collected manually by roadside enumerators in all cases until 2011, and in about half of cases in 2013. It was therefore not possible to directly check the registration marks that had been recorded in the survey for these vehicles, although registration marks which matched to vehicles that were not motorcycles in the vehicle register were assumed to be misreads. As such, these data may contain a number of misreads which would be likely to have an inflationary effect on evasion. This inflationary effect may have been significant for motorcycles, but due to the relatively small number of registration marks that were collected in this way, it is very unlikely to have had a significant effect on the overall estimates for all vehicles.

Treatment of misread registration marks in previous surveys

Roadside surveys were carried out before 2007, but in these the majority of registration marks

were collected manually and so could not be checked for accuracy, as no photographic images were available to do this. As a result, all evasion estimates given in surveys prior to 2007 did not allow for the effect of misreads, and so are likely to have been inflated and should therefore not be compared with those from subsequent surveys.

6. Strengths and weaknesses of the data

The VED estimates are based on a purpose designed statistical survey, using methods which have been carefully developed and peer-reviewed. They are based on a very large sample of more than 1 million observed vehicles per survey. The effects of numberplate misreads have been considered and additional quality assurance procedures introduced to deal with them. They therefore avoid many of the statistical biases that would be likely to arise from using data derived from administrative or enforcement systems.

However, the high cost and practical constraints of data collection mean that only relatively infrequent 'snapshots' can be taken. Surveying is concentrated in June to maximise the hours of daylight available, but this means no analysis of seasonal variation can be undertaken. The number of sites at which surveying can be carried out is also limited by cost and equipment constraints (256 sites nationally), which reduces the efficiency of the very large overall sample. The physical characteristics of motorcycles, in addition to their relatively small numbers, make them more difficult and more costly to survey than other vehicle types.

7. Uses of the roadside survey data

The estimates of VED evasion derived from this survey are published as VED Evasion Statistics. These statistics are produced primarily to provide the Department for Transport, the motoring agencies and others with evidence to take informed view on road tax evasion and related policies. The data also provide a key source of evidence for auditing the Vehicle Excise Duty account, which is independent of the processes of collecting or enforcing payment.

The same databases of observed registration marks have also been used to derive estimates of foreign vehicle traffic and MOT non-compliance, both published as DfT statistics.

Anonymised versions of the roadside survey observations, without registration marks, but with other general vehicle characteristics, are published as open data sets at www.data.gov.uk.

We welcome any feedback on these statistics, to ensure future releases best meet user needs. Feedback can be provided by email to vehicles.stats@dft.gsi.gov.uk.