Vehicle Licensing Statistics

Statistical Release

13 June 2013

Department for Transport

Experimental Statistics

Analysis of vehicle odometer readings recorded at MOT tests

This release gives an outline of the statistics which could be obtained through analysis of administrative data from the MOT testing scheme run by the Vehicle & Operator Services Agency (VOSA). It specifically focuses on analysis of vehicle mileages, derived from the odometer readings which are recorded at most MOT tests.

As a first step this note – which includes some 'work in progress' experimental statistics – is being published as a supplement to the regular Vehicle Licensing Statistics series. Our intention is to develop the statistics further in the light of further research and user feedback.

We would therefore very much welcome comments on this note to help inform and guide our work with the MOT dataset. In the future, and subject to the availability of resources, we hope to publish useful and robust statistics from it, to add to our existing portfolio of official statistics on vehicles registered in Great Britain.

A feedback form is available from our website via the Vehicle Licensing Statistics technical page, or comments can be e-mailed to vehicles.stats@dft.gsi.gov.uk.

RESPONSIBLE STATISTICIAN:
Jeremy Grove

FURTHER INFORMATION

Media Enquiries:
020 7944 3066
Public Enquiries:
020 7944 6142
vehicles.stats@dft.gsi.gov.uk

Introduction

The vast majority of cars, other light vehicles (including light goods vehicles), private buses and motorcycles in Great Britain over a prescribed age (usually three years, but in some cases one year) are subject to an MOT test. This is an annual check to ensure that the vehicle is compliant with key roadworthiness and environmental requirements. A description of the requirements for different vehicle types (test classes) is shown in the Definitions section below. Heavy goods vehicles (over 3.5 tonnes) and large buses and coaches are subject to a separate testing regime and are outside the scope of the data sets discussed in this report.

The MOT test is conducted principally at private garages and by some local authorities (for some taxis and emergency vehicles). These 'Vehicle Testing Stations' are authorised, or designated, by the Vehicle & Operator Services Agency (VOSA), an executive agency of the Department for Transport.

In 2005, VOSA introduced a computerised system for administering MOT tests and recording test results. Each vehicle testing station has secure access to a central database, MOTComp, on which the test details and outcomes are logged. This database also contains some vehicle details provided by the Driver and Vehicle Licensing Agency (DVLA) which are needed for the administration of the system. Around 38 million test records (including test passes and test failures) relating to some 27 million vehicles are stored in the database each year. Anonymised extracts of these data are available to download here. A more detailed description of the full MOT data set is published alongside these. Summary tables showing aggregate annual figures on tests and outcomes are also published as official statistics in the annual Transport Statistics Great Britain compendium.

One of the items of information recorded by the Nominated Tester during an MOT test is the vehicle odometer (milometer) reading. Although this is not currently mandatory it is recorded in around 95 per cent of cases. The MOT database therefore contains mileage information for a very large proportion of the GB-registered vehicles that are subject to an annual test. The purpose of this note is to present initial 'work in progress' estimates of vehicle mileage from the MOT data.

Total vehicle mileage at the MOT test

As MOT tests are required when a vehicle reaches 3 years of age (or in some cases one year), and annually thereafter, tests tend to take place at roughly yearly intervals, and cluster around the anniversary of a vehicle's first registration, as shown in Figure 1. This clustering means that it is possible to round the age of the vehicle to the nearest whole year for ease of analysis, without much loss of accuracy.

Figure 1: Age of vehicles in days at MOT test

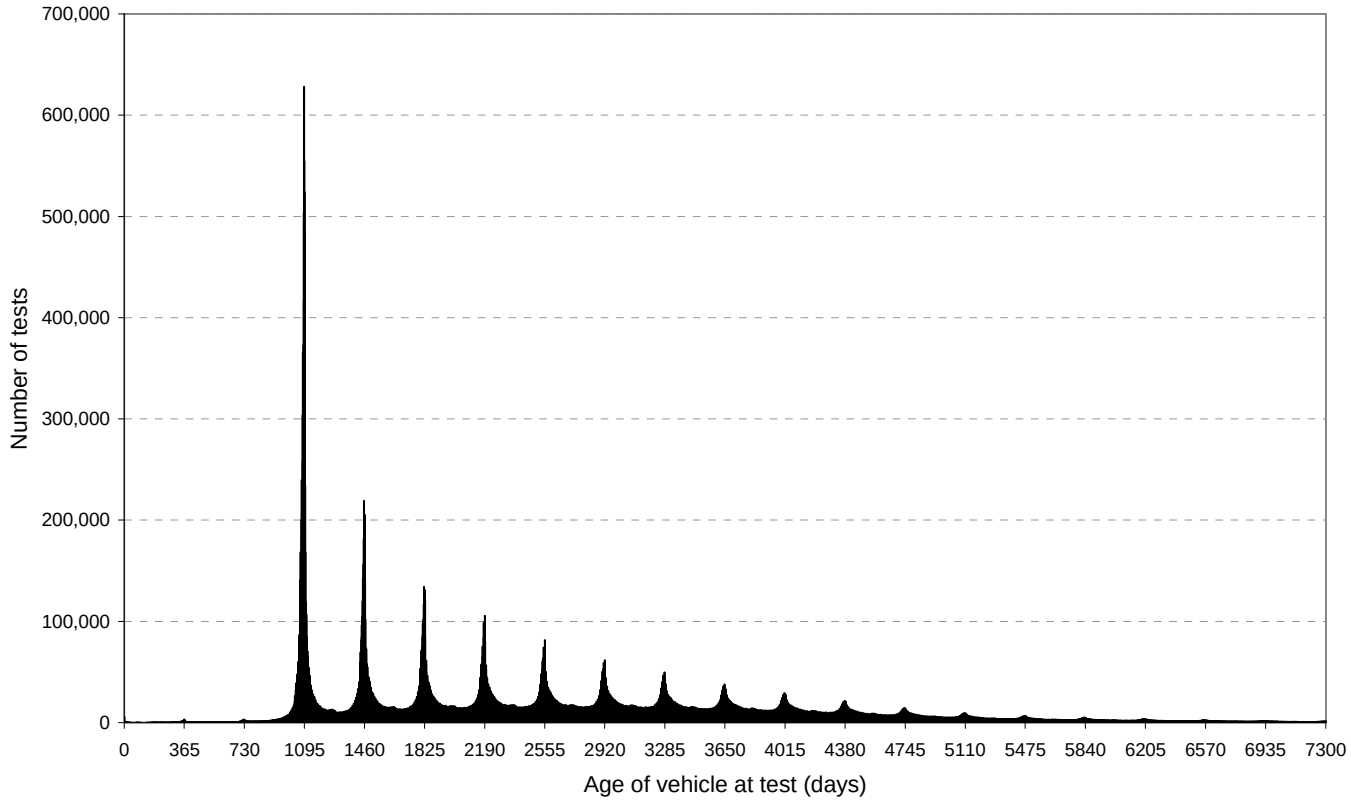


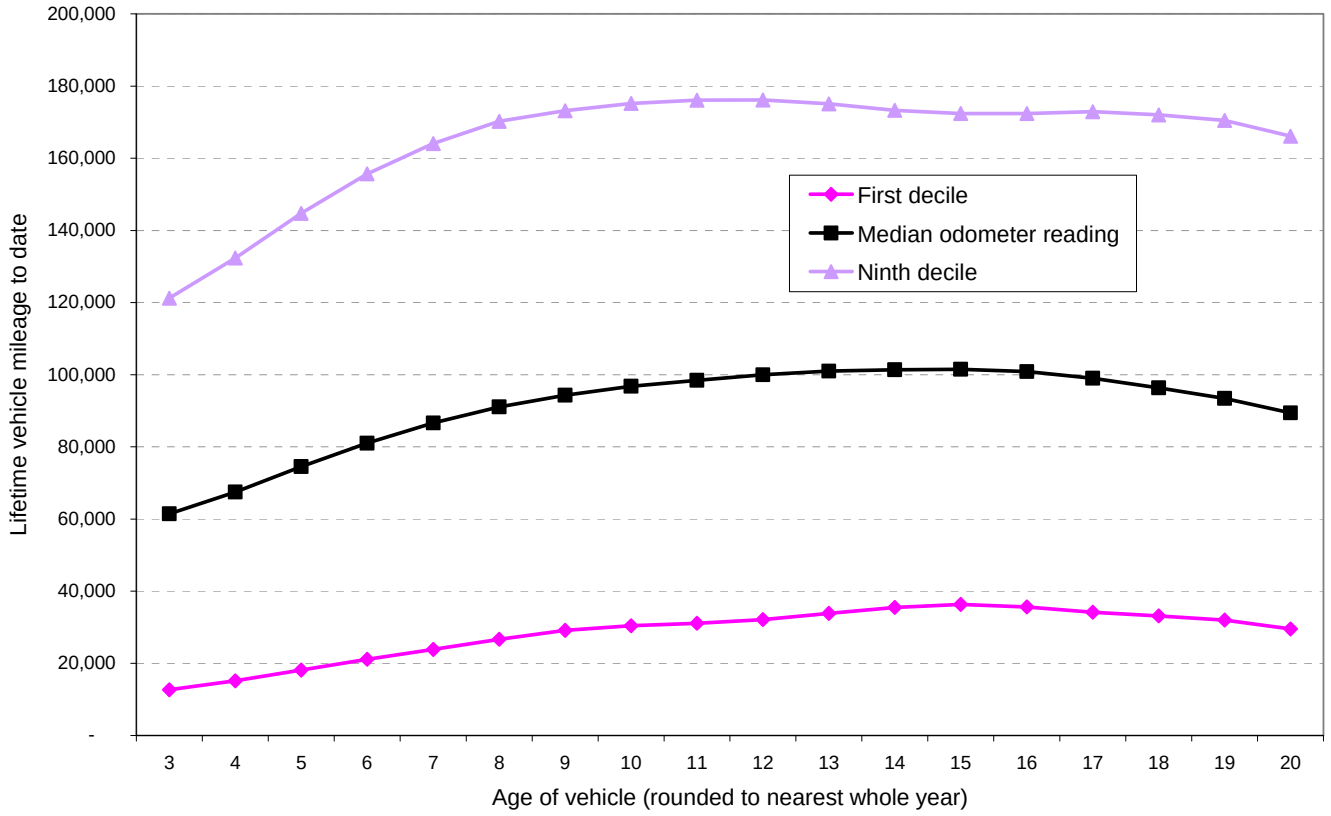
Table VEH0901 shows the distribution of raw odometer readings recorded in a subset of car MOT tests carried out in 2012, by age of the vehicle (to the nearest year) at the time of test. The raw odometer reading should represent the total mileage of a vehicle over its lifetime to that point. The median odometer reading, and a decile-based distribution of the readings, are shown in the table and in the accompanying graph (Figure 2). At this stage, we believe the median and decile points to be more reliable summary statistics for total mileage than the arithmetic mean, as they are at less risk of distortion from extreme-value data errors.

The table, and Figure 2, show a result which is at first glance counter-intuitive. While the median odometer readings of vehicles increase with vehicle age at first, the rate of increase tails off, and then actually declines with vehicle age. Clearly this is impossible for any individual vehicle but, while the results are from only a sample of the total data set, this seems to reflect a pattern across the vehicle stock as a whole of higher mileage vehicles being withdrawn from use at a younger age than lower mileage vehicles. Such vehicles would have been more heavily used and more exposed to wear and tear, and as such would be likely to reach the end of their useful life more quickly, all other factors being equal. By definition, the dataset only covers those vehicles which survive on the road to remain subject to MOT testing, and those which survive longer will tend to be those with lower mileage and less resulting wear and tear.

As vehicle age increases, the population of surviving vehicles that remain subject to MOT testing accounts for an ever-decreasing subset of the original vehicle cohort. This issue affects vehicles of all ages, but within the car fleet, there is a particularly high ‘dropout rate’ for cars in their second decade of use. In 2010, the number of MOT tests on 10-year-old cars was more than four times the number on 15-year old cars, and more than twenty times the number on 20-year-old cars.

These oldest vehicles in the dataset (those over 20 years old) therefore constitute an extremely small – and almost certainly highly atypical – subset of their original cohort

Figure 2: Cumulative mileage statistics at MOT test by age of car: MOT tests in 2012



Annual vehicle mileage between MOT tests

An annual mileage figure for all vehicles subject to MOT testing would clearly be a very useful basis for statistical analysis.

It is now possible to produce vehicle-level estimates of mileage over discrete periods by linking the successive MOT tests of a vehicle in order, and comparing the odometer readings taken at each test. Since the exact dates of tests are recorded, it is possible to refine this general approach and estimate an average annual mileage figure for the vehicle, by adjusting the difference in odometer readings between tests to an annualised figure. The approach we have taken to derive this figure has been to match consecutive MOT tests together at vehicle level to give a set of 'paired tests', usually around a year apart. The method is set out in more detail below.

The results of this test matching are not perfect, because the process relies on accurate identification of the vehicle and although the data set contains vehicle data from the DVLA vehicle register, MOT testers are also able to enter vehicle details manually. Any mistakes in manual entry of a vehicle identification number (VIN) or vehicle registration mark (VRM) can result in a failure to pair tests correctly. A second difficulty is caused by 'cherished transfers', by which a vehicle owner

may transfer a VRM from one vehicle to another. Although only a small proportion of pairings are affected by these issues, they underline the need for the data cleaning steps outlined below.

1. Create ‘pairs’ of consecutive tests of a vehicle using VIN (or alternatively VRM) to identify the vehicle, and the test count index variable within the database to identify the test number.

2. Exclude MOT tests with a ‘Pass after Rectification at the Station’ (PRS) result. This indicates that a vehicle initially failed its test, was then repaired at the test station within one hour of the test, re-tested and passed. In these circumstances two test records are generated in the data set - one fail and one PRS. Discarding the second test record (i.e. the PRS) means that the pairing of genuinely ‘separate’ MOT tests is not compromised.

3. Derive estimated annual mileage for the vehicle from the ‘paired tests’ at years *n* and *n+1* as follows:

   Estimated annual mileage associated with test n+1

   $= (\text{odometer reading}_{n+1} - \text{odometer reading}_{n}) * (365.25 / (\text{date}_{n+1} - \text{date}_{n}))$

4. Clean results by applying the following filters at vehicle level before undertaking any aggregate analysis. The underlying aim here is to remove vehicles whose annual mileage estimate is implausible, incorrect or highly atypical.

- exclude implausibly high annual mileages (over 100,000). The most likely cause of these is mis-keying of the odometer reading by the Nominated Tester. A threshold of 100,000 probably allows in significantly more anomalous values than a lower threshold of 50,000, but the latter probably excludes a modest number of genuine values. This choice can make a noticeable difference to mean annual mileage estimates.

- exclude negative mileages. These may be due to mis-keying of the odometer reading by the Nominated Tester, although it is also possible that the vehicle’s speedometer (incorporating the odometer) has been replaced, or that the odometer may have been tampered with between the tests to make the vehicle mileage appear lower (known as “clocking”). Since November 2012 the last four odometer readings have been displayed on the MOT test certificate, so the incidence of “clocking” and of data entry errors may decrease in future.

- exclude zero mileages – these are likely to represent missing data. It is arguable that some very small annual mileages (say under 250) are relatively likely to represent keying errors or mis-matched test pairs, but it is also possible that some are genuine. Excluding small values in addition to zeros can make a noticeable difference to mean annual mileage estimates.

- exclude tests conducted within 90 days of one another, as it is likely that many of these represent either:

  - vehicles which failed a MOT test and have needed remedial work before being re-tested. Such vehicles should have been kept off the road between tests, so there is likely to be little or no recorded mileage, and any recorded mileage is likely to be atypical.

  - mis-matched test pairs, which should clearly be disregarded

- exclude vehicles less than 3.5 years old at their second test. This is because most vehicles are only required to be tested once they are three years old. Some vehicles are subject to testing sooner, but for specific reasons that are likely to make them unrepresentative of the vehicle stock as a whole (e.g. cars used as taxis or private hire vehicles). As such they may have very different annual average mileages. To avoid this, only test results for vehicles 2.5 years old and over are included in these results, and since two consecutive tests are needed for a paired result, this translates to a requirement that vehicles are 3.5 or more years old at their second test.

The vehicle–level annual average mileages derived from the above process can then be analysed by a range of other vehicle and keeper characteristics by reference to the DVLA record. However, it is important to note that any resulting analysis will only cover those vehicles that are at least 3.5 years old. Newer vehicles – which are likely to exhibit different mileage patterns – are not covered.

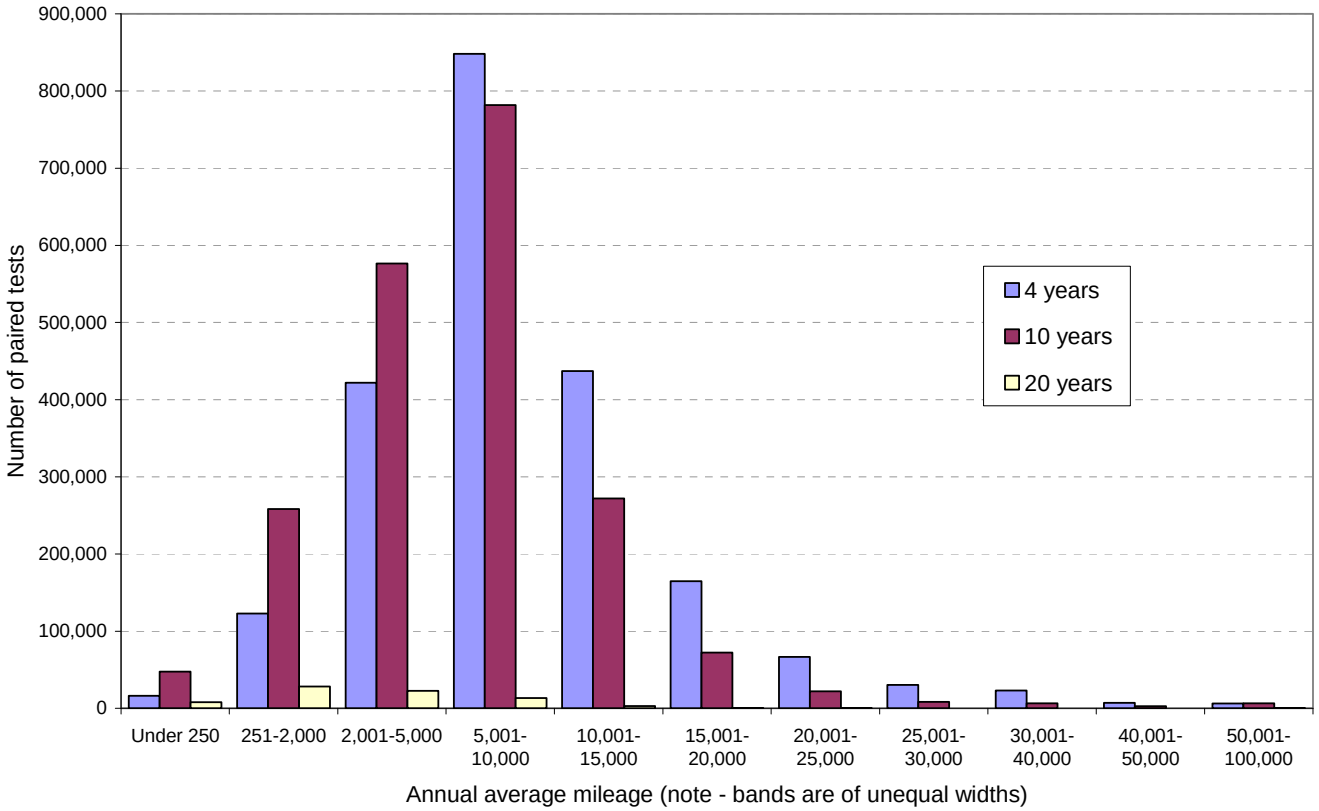
Figure 3 below shows some results for annual average estimates produced in this way. It is based on results for all cars with matched pairs of MOT tests for which the second of the tests took place in 2010. The results for three age groups of car are shown in the table – those with an age of 4, 10 and 20 years (in each case rounded to the nearest whole number of years). The number of cars falling into each annual average mileage band are shown.

This shows, first of all, that there is a wide spread of annual average mileages, especially at ages 4 and 10, and that this distribution is highly skewed – with many vehicles doing relatively modest mileages, and a fairly small number doing much more. These higher mileage vehicles will have a disproportionate effect on the mean mileage of the group as a whole.

The chart also shows that 10 year old cars are much less likely to be doing higher mileages than 4 year old cars – the distribution is shifted noticeably to the left. The greatest number of cars still fall into the 5-10,000 miles band (as is the case for 4 year old cars), but far fewer do more than 10,000 miles a year, and far more less than 5,000 miles, than for 4 year old cars.

This shift towards lower mileages is continued further for the 20 year old group – the most common mileage band for this group is the 250-2,000 miles category. But the most obvious feature of the 20 year old group is how few of them there are, relative to the 10 year old group. As noted earlier, it will tend to be the higher mileage vehicles which have been scrapped before reaching this age.

Figure 3: Annual average mileage bands of cars at ages 4, 10 and 20 years: second MOT test of pair in 2010



Mileage at first MOT test

It is also possible to estimate an annual average mileage for the first three years of a vehicle's life, by using the mileage at the date of the first MOT test. This can then be adjusted to cover an exact 3-year period by comparing the date of the first MOT test to the date of the vehicle's initial registration, similar to the 'annualisation' process described in the previous section.

Clearly, the MOT dataset does not contain any information about the profile of a vehicle's mileage over its first three years of use. However, we can use other datasets (notably the National Travel Survey) to produce a general estimate of this profile.

Further development

The results and methods outlined above represent only an introduction to the analyses of the MOT data which are possible. Some of this work is possible using the anonymised data sets which have already been published online. However, some analyses require the use of personal data fields which are not publicly released. Further DfT work is therefore likely to concentrate particularly on analyses which require these additional fields, with the aim of producing useful statistics for publication at aggregate (rather than record) level.

Possible areas for further investigation include:

- Small area statistics, based on the postcode of the registered keeper of the vehicle, and possibly the postcode of the MOT testing station
- Analysis of mileage statistics against vehicle or keeper characteristics

Alternative sources of vehicle mileage data

The department already produces two other sets of statistics relating to vehicle mileages – one from interviewing householders about how far they travel by road in the National Travel Survey, and one from measuring traffic on the roads. These data sources are designed for different purposes and so provide different coverage of vehicle mileage.

The choice of which of these statistics is the best to use is highly dependent on how it will be used. A guidance note has been published alongside this document to help users choose which data source is most suitable for them (see Vehicle Licensing Statistics technical page.):

Strengths and weaknesses of the MOT dataset as a source of statistics

The previous sections outlined in some detail the key characteristics of the vehicle mileage statistics that can be derived from the MOT dataset. In summary the key points are:

Strengths

- The data set covers a very large proportion of licensed cars, light vans and motorcycles in Great Britain (effectively almost all such vehicles more than 3 years old). This coverage much larger than any can be obtained by any survey method
- The size and coverage of the data set, combined with the additional vehicle and keeper details available, allow detailed analyses of the data to be produced for relatively small sections of the population (subject to usual controls to prevent disclosure of individual data)
- The data are readily available as a by-product of an existing administrative process and therefore have minimal additional collection cost

Weaknesses:

- Most vehicles under 3 years old are excluded, and those which are included have specific, atypical characteristics

- HGVs and most buses and coaches are tested under a different system, and are therefore also excluded
- It is dependent on the quality of information recorded at the time of the MOT test – in particular the transcribed odometer reading (a large integer, therefore relatively prone to error). This has little scope for data cleaning other than excluding implausible or unusual values
- Any overseas or off-road mileage driven by a vehicle will be included in its odometer reading and cannot be identified or discounted. Although this will be a relatively small proportion of total mileage in most cases, it does affect comparability with other DfT statistics on road usage.

The DfT would be grateful for comments on the strengths and weakness of the experimental results presented, and also on the potential for further analysis of the data source.

A feedback form is available from our website via the Vehicle Licensing Statistics series page, or comments can be e-mailed to vehicles.stats@dft.gsi.gov.uk.

Background notes

1. The data table accompanying this note (VEH0901) is available for download from the Vehicle Licensing Statistics series page.

2. The statistics published in and alongside this document are experimental, and may be further developed in the light of further research and feedback from users. They have been produced in accordance with all relevant provisions of the Code of Practice for Official Statistics, but have not been designated as National Statistics.

3. Anonymised MOT testing data sets are available, along with an MOT Testing Data User Guide, here.

4. General information about the MOT test is available here.

Definitions

Test Classes

The vehicles subject to test under the Regulations are divided into the following classes: -

Class	Description	Age at which 1st test required (years)
1	Motor bicycles (with or without sidecars) up to 200 cm^3	3
2	All motor bicycles (including Class 1) (with or without sidecars).	3
3	3 wheeled vehicles not more than 450 kg unladen weight (excluding motor bicycles with side cars). (3 wheeled vehicles more than 450 kg unladen are in class 4.)	3
4	Cars, passenger vehicles, motor caravans, Private Hire Vehicles, Motor Tricycles, Quadricycles and dual purpose vehicles in all cases with up to eight passenger seats	3
	Goods vehicles not exceeding 3,000 kg Design Gross Weight (DGW).	3
	Taxis and ambulances in either case with up to eight passenger seats.	1
	Passenger vehicles, ambulances, motor caravans and dual purpose vehicles in all cases with nine to twelve passenger seats that; • are fitted with no more seat belts than the minimum required because of their construction; or • are identified as having been fitted with a type approved seat belt installation when built; or • have been tested as class 4A, 5A or 6A (PSV) with at least the same number of seat belts as are currently fitted.	1
4A	**The class 4A test is the normal class 4 test with the addition of a check on the installation of certain seat belts**. Passenger vehicles, ambulances, motor caravans and dual purpose vehicles in all cases with nine to twelve passenger seats that: • are fitted with more seat belts than the minimum required because of their construction and: • are not identified as having been fitted with a type approved seat belt installation when built; or • have not been tested as class 4A, 5A or 6A (PSV) with at least the same number of seat belts as are currently fitted.	1
5	Private passenger vehicles, ambulances, motor caravans and dual purpose vehicles in all cases with thirteen or more passenger seats (including community and play buses, etc.) that: • are fitted with no more seat belts than the minimum required because of their construction; or • are identified as having been fitted with a type approved seat belt installation to all seats when built; or • have been tested as class 5A or class 6A (PSV) with at least the same number of seat belts as are currently fitted.	1
5A	**The class 5A test is the normal class 5 test with the addition of a check on the installation of certain seat belts**. Passenger vehicles, ambulances, motor caravans and dual purpose vehicles in all cases with thirteen or more passenger seats (including community buses, etc.) that: • are fitted with more seat belts than the minimum required because of their construction and: • are not identified as having been fitted with a type approved seat belt installation when built; or • have not been tested as class 5A or class 6A (PSV) with at least the same number of seat belts as are currently fitted.	1
7	Goods Vehicles over 3,000 kg up to and including 3,500 kg DGW	3