Analyses of driver licence records from DVLA

by J Broughton

TRL Report 403

E

ISSN 0968-4107

Analyses of driver licence records from DVLA

TRL REPORT 403

TRANSPORT RESEARCH LABORATORY

Analyses of driver licence records from DVLA

Prepared for Road Safety Division, Department of the Environment, Transport and the Regions

J Broughton

TRL REPORT 403

First Published 1999
ISSN 0968-4107

This report has been produced by the Transport Research Laboratory, under/as part of a Contract placed by the Department of the Environment, Transport and the Regions. Any views expressed are not necessarily those of the Department.

TRL is committed to optimising energy efficiency, reducing waste and promoting recycling and re-use. In support of these environmental goals, this report has been printed on recycled paper, comprising 100% post-consumer waste, manufactured using a TCF (totally chlorine free) process.

Transport Research Foundation Group of Companies

Transport Research Foundation (a company limited by guarantee) trading as Transport Research Laboratory. Registered in England, Number 3011746.

TRL Limited. Registered in England, Number 3142272.
Registered Offices: Old Wokingham Road, Crowthorne, Berkshire, RG45 6AU.

CONTENTS

Executive Summary

A previous report (Broughton, 1986) described the data held in the Driver Licence File of the Driver and Vehicle Licensing Agency (DVLA) at Swansea. The potential for analysing these data as part of road safety research was discussed and several analyses of the data available at that time were presented. The new report presents the results of a recent series of analyses of the data from this file.

The rate at which drivers are convicted of driving offences is one important area that was analysed in the earlier report. The timescale of these analyses was limited, however, by the relatively short period after which details of driving offences may be removed from licence records, in accordance with the relevant legislation. A driver may apply for a 'clean' licence four years after being convicted of an offence, at which time the relevant details are removed from the computer file (the minimum period is eleven years for some more serious categories of offence). This, combined with the slowness with which some courts supplied information to DVLA, meant that many analyses were restricted to offences committed in a single calendar year, 1983.

An archive was set up at TRL in order to overcome this loss of information: it is maintained purely for statistical and scientific purposes, and has been registered in accordance with the Data Protection Act, 1985. Three times a year, selected details from a sample of DVLA driver licence records are sent to TRL, and the archive is updated whenever new or changed information is detected. Thus, the TRL file contains all of the current DVLA data of interest for the sample of drivers, *plus* any details which have been removed from the DVLA file. The sampling system ensures that a representative group of new drivers is automatically added to the TRL file each time that the archive is updated. The archive thus allows the details of the offences which have been committed over more than a decade to be linked to the offenders' personal details such as age, sex and region of residence.

The report presents a series of analyses which focus upon the rate of offences per thousand drivers per year over the period 1985-97 to see how the rate varies with age and sex of the driver and region of the country. The DVLA file only contains details of those driving offences which are sufficiently serious to require the offender's driving licence to be endorsed, so lesser offences such as parking and obstruction are not included.

The overall rate of driving offences is found to vary considerably by age and sex of driver. Men commit many more offences than women, and the rate tends to fall with age. The overall offence rate rose in the late 1980s among men with full driving licences, but has since changed little. The rate among men with provisional licences, however, has fallen steadily in the 1990s and is now well below the rate among men with full licences. The chief factor contributing to this fall has been the rapid reduction in the rate of Licensing offences.

Three groups of offences have also been studied: Speeding, Careless Driving and Drink/Driving. Different patterns were found in each case. The rate of Speeding offences rose through most of the period studied among drivers with full licences, but fell among drivers with provisional licences; in any particular year, the rate fell gradually with age. The rate of Careless Driving offences fell through most of the period studied, and in any single year fell rapidly with age. The rate of Drink/Driving offences among male drivers with full licences fell until 1993 but has since risen slightly; the rate has fallen in most years among male drivers with provisional licences and has remained low among female drivers. In any single year, the rate of these offences among male drivers with full licences is greatest in the 20-24 age group, but in the 25-29 age group for those with provisional licences.

Offence rates are also examined by region, looking both at average rates for the 1993-97 period and changes since the 1987-92 period. The highest 1993-97 rates were found in North West England, for both male and female drivers; several regions have rates well below those in the North West, the lowest rate for male drivers was in London and for females the lowest rate was in the East Midlands. Male rates appear to have fallen least in North West England and most in the West Midlands; female rates appear to have risen most in South West England and fallen most in the West Midlands.

The relatively small number of offences in the TRL archive from some regions limits the comparisons that can usefully be made of regional rates for specific types of offence. It does appear, however, that the rate of Speeding offences is highest in North West England, almost twice as high as in London. A similar range was found with Careless Driving offences by men, but in this case the lowest rate was in South East England. The range was much less with Drink/Driving offences.

Many drivers commit no offences, a few commit many offences. Among drivers who have held a full licence for more than five years, 82.8 per cent of men and 93.9 percent of women had committed no offences in the past five years. 0.04 per cent of men had committed at least 10 offences, and effectively no women. However, when all drivers are considered (i.e. drivers with a provisional licence or no licence are also included), the percentage rises to 0.12 for men and 0.01 for women.

Reference

Broughton J (1986). *Analysis of motoring offence details from DVLC driving licence records*. Research Report RR77. Transport Research Laboratory, Crowthorne.

1 Introduction

An earlier report (Broughton, 1986) described the data held in the Driver Licence File of the Driver and Vehicle Licensing Agency (DVLA) at Swansea. It discussed the potential for analysing these data as part of road safety research, and presented a number of analyses of the data that were available at that time. This report presents the results of a recent series of analyses which has been carried out of data from this file.

The rate at which drivers are convicted of driving offences is one important area that was analysed in the earlier report. The timescale of these analyses was, however, limited by the relatively short period after which details of driving offences may be removed from licence records, in accordance with the relevant legislation. A driver may apply for a 'clean' licence four years after committing an offence, at which time the relevant details are removed from the computer file (the minimum period is eleven years for some more serious categories of offence). This, combined with the slowness with which some courts supplied information to DVLA, meant that many analyses were restricted to offences committed in a single calendar year, 1983.

An archive was set up at TRL in order to overcome this loss of information: it is maintained purely for statistical and scientific purposes, and has been registered in accordance with the Data Protection Act, 1985. Three times a year, selected details from a sample of DVLA driver licence records are sent to TRL, and the archive is updated when new or changed information is detected. Thus, the TRL file contains all of the current DVLA data of interest for the sample of drivers, *plus* any details which have been removed from the DVLA file. The sampling system ensures that a representative group of new drivers is automatically added to the TRL file each time that the archive is updated. The archive allows the details of the offences which have been committed over more than a decade to be related to the offenders' personal details such as age, sex and region of residence. Naturally, this is only possible within the range of information that drivers supply to DVLA.

The chief source of published information about driving offences is an annual series of Home Office Statistical Bulletins (e.g. Wilkins and Addicott, 1998). These bulletins provide aggregate details of the number and type of offences committed by the population of England and Wales, together with details of the court system. They do not link offences to the details of the offenders, however, so the analyses presented in this report complement the statistics published by the Home Office.

Section 2 describes the data from DVLA that are held in the TRL archive. Section 3 then presents a series of analyses of the data relating to driving offences, based on the data that were available in April 1998; the analyses cover the period 1985-1997. Section 4 examines trends in the number of licensed drivers and Section 5 brings together the conclusions that may be drawn.

2 Driving licence information

The DVLA computing facilities have been extensively modernised in the decade since the previous report, but the organisation of the Driver Licence file has not changed greatly from the state described there. The file is ordered by the driver number (which appears on any driving licence); this is based on the first five letters of the driver's surname and a six digit code derived from the date of birth. The effect is that the file consists of a sequence of groups of records, with one group containing the records of drivers whose surnames share the first five letters. Within each group, drivers born in the same decade are brought together, with male drivers preceding female drivers.

The sample of drivers represented in the TRL archive consists of all those with surnames in the range CHAME-CHEND and SWEET-TAYLL. These ranges were chosen when the archive was established in 1986 with the aim of avoiding bias due to regional surnames: each range contains approximately 0.6 per cent of all driver licence records. The earlier report found that there was some regional bias, in particular a shortage of Scottish drivers. This may well have been avoided with an alternative sampling strategy, such as choosing every N-th record, but the strategy adopted has two important advantages:

- it is simple and cheap to implement;
- it automatically includes a sample of all those applying for their first driving licence.

The task of maintaining the TRL archive of DVLA licence data for this sample of drivers is conceptually simple but complex in detail. Every four months, a computer file is prepared at DVLA which contains the current values of a certain range of items for all drivers in the two name ranges. These values are compared with the values in the TRL file and a new file is created which combines the new (DVLA) and existing (TRL) data. Thus, new data are added to the TRL file, but existing data are rarely modified: the main reason for modification is the removal of an offence record from the DVLA file following a successful appeal.

In the earlier report, analyses for the two name ranges were carried out in parallel. The results differed in detail, but the differences were within the range that might be expected with two relatively small samples of the full population of drivers. Accordingly, the results presented below are based on analyses of the combined name ranges.

2.1 Driving licence details

The following is a summary of the data held on the TRL file. Not all data are present for all records: for example, there are no details of driving tests for drivers who have passed no test or who passed before the DVLA system became fully operational in the late 1970s. (In the latter case, the existence of an entitlement to drive implies that a test was passed or, for elderly drivers, that they were driving before the driving test was introduced). Data include:

i personal details such as surname, initials, title, date of birth, postcode and groups of vehicle for which an entitlement exists;

ii date of expiry of licence;

iii date of issue of first licence;

iv details of driving test passes;

v medical markers indicating whether the driver's declared medical condition may affect his ability to drive;

vi motorcycle provisional licence details;

vii details of convictions for driving offences.

Details of convictions for driving offences are supplied by courts after the cases have been tried, so the DVLA records are sometimes updated several months after the offence was committed. The date of each offence and conviction is recorded, together with a four character offence code taken from the DVLA system of 'Endorsement offence codes'. The first two characters of an offence code indicates the general type of offence, while the remaining digits are more specific: for example, AC denotes 'Accident offences' while AC10 denotes 'Failing to stop after an accident'. Offences which do not involve penalty points are not endorsable and so do not appear in the DVLA records. Thus, all analyses reported below relate to *endorsable* offences. Lesser offences, such as parking and obstruction offences, are excluded and do not contribute to the offence totals.

The groups of offence codes are:

AC	Accident Offences
BA	Driving while Disqualified
CD	Careless Driving
CU	Construction & Use Offences
DD	Reckless/Dangerous Driving
DR	Drink or Drugs
IN	Insurance Offences
LC	Licence Offences
MS	Miscellaneous Offences
MW	Motorway Offences
PC	Pedestrian Crossings
PL	Provisional Licence Offences
SP	Speed Limits
TS	Traffic Direction and Signs
UT	Theft or Unauthorised Taking

The system is periodically adjusted to reflect changes in the driving laws. Most UT offences ceased to be endorsable in July 1992 and there were lesser changes such as the replacement of BA20 (Driving while disqualified as under age) by LC20 (Driving otherwise than in accordance with a licence) from the same date. Such changes will have a minor influence on the changes with time in the number of offences recorded. Other factors such as police enforcement policy may also have some influence. For example, the option of fixed penalty notices for endorsable offences such as speeding and traffic light violations (introduced in October 1986) may well have increased enforcement.

2.2 Changes of surname

The TRL archive suffers from one inescapable problem: a small number of drivers notify DVLA each month of a changed surname. Marriage is the main reason for surnames to change, so the effect is more pronounced for women than for men. A secondary reason is the correction of names that had been entered wrongly in the DVLA file. There are three cases:

a when the old name is within one of the ranges but the new name is not then no current data can be added to the record of this driver, so the records of drivers who 'leave' the archive cannot be updated,

b conversely, when the old name is not within one of the ranges but the new name is then some earlier data may be missing from the record of this driver, i.e. the information for drivers who 'join' the archive may be incomplete,

c when both old and new names are within one of the ranges, the earlier data for the old name can be transferred to the record for the new name, so there is no loss of information.

Corrections are generally minor, so tend to fall under (c); however, most changes following marriage fall under (a) or (b). The only remedy is to restrict analyses to drivers whose names have not changed during the period being studied, so in effect the archive contains proportionately fewer women drivers than men. To ensure that the results presented in this report include only driver records which are complete, all analyses will exclude any driver who either:

left the archive before the end of the period studied, or

joined the archive less than two years after the beginning of the period studied (since details for the interim period should have been complete at the time of joining and have since been preserved in the archive).

In order to explore the loss of drivers due to name-changes, Table 1 displays the Sampling Fraction at two dates. This is defined as:

$$\text{Sampling Fraction} = \frac{\text{number of licensed drivers in TRL archive}}{\text{number of licensed drivers in DVLA file}}$$

where the number of licensed drivers in the DVLA file comes from counts of the actual file made by DVLA at the two dates. Two sets of fractions were calculated from the TRL file as it existed in April 1998:

i to prepare the first row, the number of licence-holders at the earlier date was reconstructed from licensing details showing when provisional and full licences were issued, then divided by the number shown by the DVLA count at that date;

ii to prepare the second row, the number of licence-holders was calculated directly when the TRL file was updated, then divided by the number shown by the DVLA count at this date.

Table 1 Sampling fraction at two dates calculated from the TRL file in April 1998

	Provisional licences		*Full licences*	
	men	*women*	*men*	*women*
28 December 1989	1.25%	0.98%	1.12%	0.96%
5 April 1998	1.28%	1.18%	1.15%	1.12%

The four percentages in the lower row have suffered no loss due to name-changes and show the relation between the TRL and DVLA files at the point when the TRL file was updated. Analyses made when the TRL file was updated previously have found identical percentages, so the number of records in the two name ranges has grown exactly in step with the growth of the DVLA file.

The percentages in the upper row are based on those drivers who *were* in the TRL file in 1989 and still were in 1998, so if there had been no name-changes then the two rows would be identical. Thus, there has been scarcely any loss of information due to name-changes among men. Among women, however, the sampling fractions are one-seventh lower for the earlier date, showing that the number of records has fallen by this proportion due to name-changes over this period. Comparisons by age show that the fall is indeed concentrated among younger women, confirming that the great majority of name-changes notified to DVLA are the result of marriage.

2.3 Calculation of offence rates

The loss of drivers due to name-changes should not bias the results achieved, provided that it is taken into account when designing analyses. The main way to achieve this is calculate rates, linking the number of offences recorded in the file to the 'exposure' of drivers to the possibility of committing an offence. This approach is followed wherever possible in the analyses presented below.

The measure of exposure would ideally include factors which might influence the number of offence committed, such as drivers' annual mileages, but these are not included among the DVLA data. The best available measure is the actual number of drivers in the file with a particular type of licence at that time. For example, the rate of driving offences in the period 1990-91 committed by 20-29 year old men with full driving licences is calculated by:

1 considering all offences in the file that were committed in 1990-91, and checking whether at the time of the offence the offender was:
   a a 20-29 year old man (based on his personal details) who
   b had a full driving licence (based on the licence details and the date of any test pass);
2 checking how much of the period 1990-91 each man in the file spent aged 20-29 with a full driving licence, based on his date of birth and of the date of any test pass, and accumulating the individual results to give the overall 'exposure' of this group of licence-holders;
3 dividing the number of offences from 1. by the number of driver-years from 2. to obtain the rate of offences per driver-year.

Drivers who have left the archive as the result of name-changes will be excluded from the calculation, which reduces the number of offences as well as the total exposure; provided that the age ranges are reasonably narrow, both should be reduced to the same degree so that the rate should not be affected significantly. Appendix A shows how rates can be adjusted to avoid bias when wider age ranges are used. The adjustment is only made for women, and even here the effect is only minor: for example, it only raises the overall offence rate for women in 1988-97 calculated in Section 3.1 by 4.0 per cent.

The British driver licensing regulations are highly detailed, so an individual's driving licence can develop in many different ways, for example:

- many drivers are entitled to drive several categories of vehicle, so each of these drivers progressed from provisional to full entitlement at several dates;
- a provisional licence with motorcycle entitlement is withdrawn after two years if the rider has not passed a test;
- some drivers are disqualified from driving following conviction for a driving offence.

Nevertheless, the majority of licences follow a common pattern, and the following definition has been adopted in order to simplify the analyses and the presentation of results which relate to the status of a licence:

i a driver is unlicensed until receiving a provisional licence;
ii a driver then has a provisional licence until passing a driving test;
iii a driver then has a full licence until the expiry of that licence as shown in the current file;
iv a driver is then unlicensed.

When a driver is disqualified following conviction for a driving offence, the driving licence is removed and he or she becomes unlicensed. The analyses take account of periods of disqualification, so at any particular time the category 'unlicensed driver' includes not only those who had yet to acquire a licence and those whose licence had expired but also those who were disqualified from driving at that time.

3 Driving offences 1985–97

As explained in the Introduction, the principal reason for setting up the archive system was to preserve details of driving offences which are lost from the DVLA file. This allows larger numbers of offences to be studied, giving rise to more precise results, and it also allows developments over time to be analysed.

The type of vehicle that was being driven or ridden at the time of an offence cannot be identified from the data held in the DVLA file. Consequently, it is impossible to investigate the relative rates of car drivers, lorry drivers, motorcyclists etc. Equally, the file contains no information about other factors that might affect the number of offences, such as the annual mileage.

Drivers are sometimes prosecuted for several offences on a single occasion. For example, the police may stop a driver for careless driving and subsequently find that the driver was not insured, or that the vehicle was defective. Several tables refer to 'incidents', where all offences committed on a single day are counted as one.

Sections 3.1 and 3.2 present various *direct* analyses of the offence data between 1988 and 1997 which avoid the use of

statistical models. This approach has the advantage that the method can be readily understood by non-statisticians, but there are also certain disadvantages, namely:

a changes over time cannot be investigated properly; the decade could be divided into a number of shorter periods which could then be compared, but there would be correspondingly fewer offences per period so apparent differences between periods could simply be the result of chance;

b some groups of driver committed relatively few offences, so comparisons with other groups could be affected by the effects of chance: this is especially important since only a relatively small sample of driver licence records is available for analysis.

Accordingly, Section 3.3 presents *indirect* analyses which use an appropriate statistical model. Section 3.4 uses the same type of model to compare offence rates by region.

3.1 Direct analyses of offence rates

The first and most direct analysis of the offence data is presented in Figure 1, to provide an overview of the data. The analysis focuses on the period 1993-97 and, for each offence committed during these years by the licence-holders in the sample, checks their age at the time of the offence. The Figure shows the number of offences committed by year of age, without considering the numbers of licence-holders and the types of licence held. The annual average numbers are shown for the more frequent offence types, using a logarithmic scale and slightly smoothing the graphs for the less frequent types.

The main points to emerge are that men commit far more offences than women, and that young drivers commit far more offences than older drivers; however, the pattern of decline with age differs with the type of offence. The peak age for offending (all types of offence) is 23 among men and 28 among women. Speeding (SP) is the most frequent type of offence for women of all ages and men from 20: the peak occurs at 28 among women and 31 among men. The graphs for most other types of offence follow a different pattern, with peaks at or before the age of 20 followed by declines; however, Drink/driving (DR) and Traffic Direction (TS) offences peak much later, around the age of 30.

This preliminary analysis does not consider the 'exposure' of drivers as defined in Section 2.3, so the results could be influenced by the distribution of licence-holders by age and sex. The next analysis calculates the number of offences per thousand driver-years during 1988-97 by following each driver in turn, to see when they gain a provisional and then a full licence, and relating any offences committed to the licence held and their age at the time. This gives the number of offences as well as the exposure - i.e. the number of driver-years during the period studied. The rates calculated over all ages are shown in Table 2, together with exposure: the latter are included to indicate the relative precision of the offence rates, for rates based on one million driver-years of data are less vulnerable to sampling effects than those based on one thousand. Exposure and rates for women have been adjusted to compensate for name-changes, as described in Appendix A, to be directly comparable with the rates for men.

Table 2 Offences per thousand driver-years, 1988–97

	Men with a		*Women with a*	
Type of offence	*Full licence*	*Provis -ional licence*	*Full licence*	*Provis -ional licence*
Accident Offences	0.9	1.2	0.2	0.2
Careless Driving	3.1	2.1	0.9	0.3
Construction &Use Offences[1]	2.2	2.9	0.2	0.2
Reckless/Dangerous Driving	0.2	0.5	0.0	0.0
Drink or Drugs	3.9	4.7	0.5	0.3
Insurance Offences	4.4	23.4	0.7	1.5
Licence Offences[2]	0.4	14.5	0.1	1.2
Provisional Licence Offences[2]	0.1	15.3	0.0	1.0
Miscellaneous Offences	0.1	0.1	0.0	0.0
Motorway Offences	0.4	0.5	0.0	0.0
Pedestrian Crossings	1.7	0.3	0.6	0.1
Speed Limits	29.1	4.2	8.8	0.3
Traffic Direction and Signs	4.9	1.1	1.5	0.1
Theft or Unauthorised Taking[3]	0.1	2.1	0.0	0.0
All offences	51.6	73.9	14.2	5.4
All incidents	50.0	51.2	13.8	3.1
Exposure (driver-years×10^3)	2027	231	1255	298

[1] certain of these offences were not endorsable from July 1992
[2] included among Licence Offences from July 1992
[3] most of these offences were not endorsable from July 1992

Tables 3 and 4 show the rates by age group, together with the rates for the more common types of offence. Figure 2 presents the overall offence rates from these tables, and shows clearly the very regular decline with age. Rates are highest in the 17-19 age group, with the minor exception of women drivers with provisional licences. On average for this age group, 1-in-6.9 men and 1-in-34 women with full licences committed a driving offence each year between 1988 and 1997, compared with 1-in-9.5 men and 1-in-156 women with provisional licences.

3.2 Further direct analyses

3.2.1 Unlicensed drivers

The rates in Tables 2-4 relate to offences committed by licensed drivers, where the licence records provide the exposure data needed to calculate offence rates. However, many offences are committed by unlicensed drivers: those who have yet to acquire a licence, whose licence has expired or who have been disqualified from driving for a period. The DVLA Driver Licence file naturally contains records of the second and third groups; it also contains records of the first group in case these individuals subsequently apply for a driving licence. Consequently, it is possible to calculate the percentage of offences that were committed by unlicensed drivers. Table 5 shows these figures for those types of offence where the number committed by unlicensed drivers accounted for more than 0.3 per cent of all offences. Reckless or Dangerous Driving is also included as the only other type of offence where more then one tenth of offenders were unlicensed. All disqualified drivers are technically unlicensed, so 100 per cent of these offences were committed by unlicensed drivers.

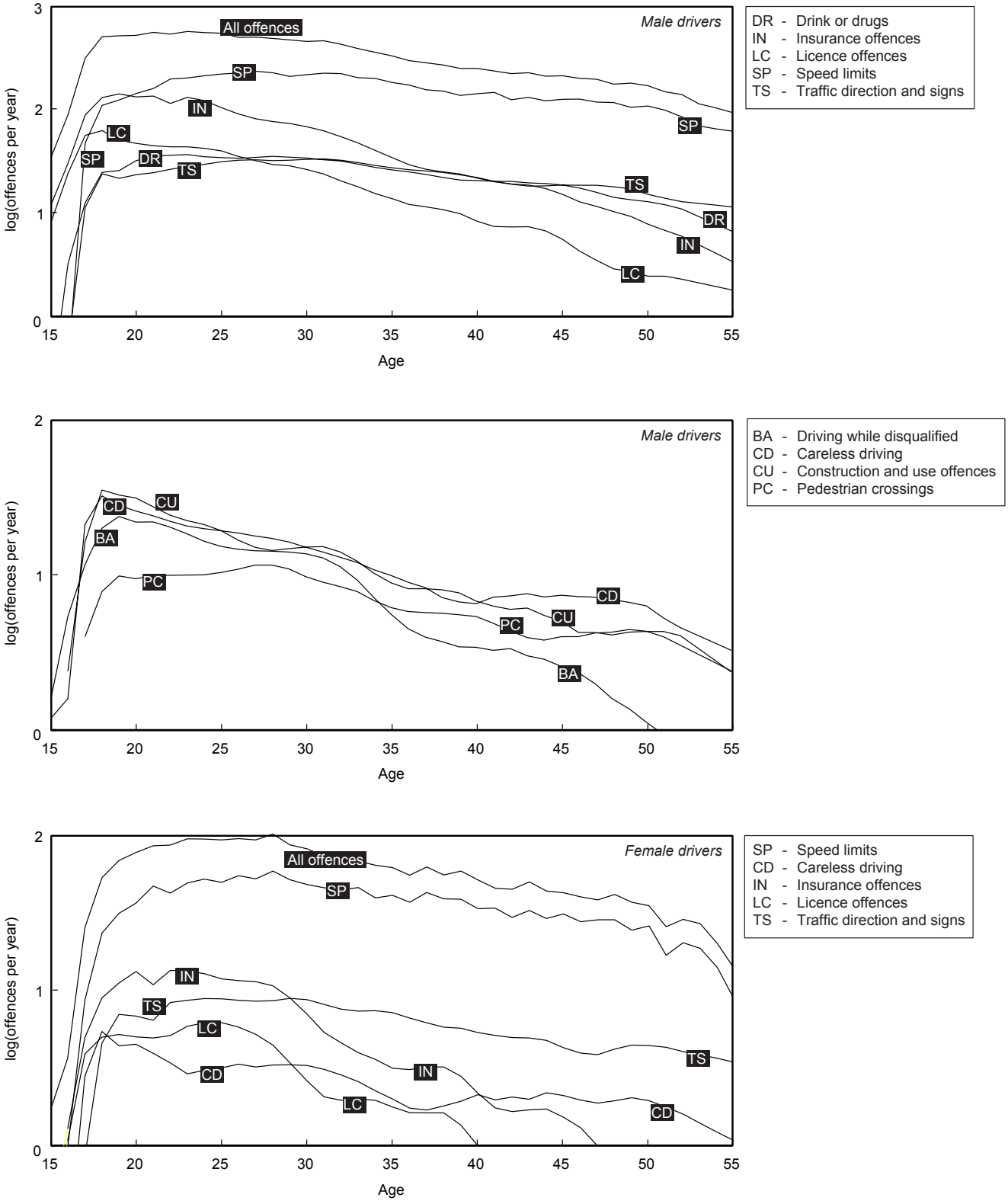


Figure 1 Annual number of offences by sample of drivers in 1993–97

Table 3 Offences per thousand driver-years, men, 1988-97

Age	*17-19*	*20-24*	*25-29*	*30-39*	*40-49*	*50-59*	*60-69*	*70-*	*All*
Full licences									
Careless Driving	16.3	7.7	4.3	2.6	2.0	1.7	1.1	1.4	3.1
Drink or Drugs	7.0	8.6	6.5	4.9	3.9	2.1	0.7	0.2	3.9
Insurance Offences	21.6	16.9	8.7	3.9	1.9	0.8	0.2	0.1	4.4
Speed Limits	57.6	55.0	48.2	36.4	27.3	18.5	7.0	2.6	29.1
All offences	144.0	115.1	86.1	59.9	43.3	29.3	12.0	5.8	51.6
All incidents	132.7	108.1	82.0	57.6	41.9	28.6	11.6	5.5	49.2
Exposure (driver-years×10^3)	57	176	221	442	404	308	268	151	2027
Provisional licences									
Careless Driving	3.3	2.4	2.0	1.2	1.5	1.4	1.0	2.8	2.1
Drink or Drugs	3.6	5.1	6.3	5.4	4.5	2.6	1.1	0.0	4.7
Insurance Offences	33.7	36.4	25.2	14.6	9.7	5.3	0.8	2.8	23.4
Speed Limits	4.1	2.8	3.7	5.0	7.4	3.4	0.4	0.0	4.2
All offences	105.5	102.0	76.4	52.3	44.0	25.4	8.5	12.5	73.9
All incidents	72.4	70.5	53.9	35.7	32.0	16.7	5.2	8.3	51.2
Exposure (driver-years×10^3)	48	47	41	50	24	13	7	1	231

Table 4 Offences per thousand driver-years, women, 1988-97

Age	*17-19*	*20-24*	*25-29*	*30-39*	*40-49*	*50-59*	*60-69*	*70-*	*All*
Full licences									
Careless Driving	3.7	1.6	1.0	0.7	0.8	0.6	0.5	0.9	0.9
Drink or Drugs	0.4	0.8	0.7	0.6	0.6	0.2	0.1	0.0	0.5
Insurance Offences	2.7	2.5	1.4	0.7	0.4	0.1	0.0	0.0	0.7
Speed Limits	16.3	15.0	13.0	10.1	8.4	5.8	2.4	1.1	8.8
All offences	29.6	25.2	19.9	15.0	12.5	8.4	3.8	2.8	14.2
All incidents	28.6	24.4	19.4	14.6	12.2	8.2	3.8	2.6	13.8
Exposure (driver-years×10^3)	41	143	185	349	297	189	125	60	1389
Provisional licences									
Careless Driving	0.3	0.4	0.4	0.2	0.1	0.2	0.1	1.1	0.3
Drink or Drugs	0.1	0.3	0.5	0.4	0.1	0.1	0.1	0.0	0.3
Insurance Offences	2.2	3.1	2.2	1.0	0.6	0.1	0.2	0.0	1.5
Speed Limits	0.2	0.3	0.3	0.5	0.4	0.1	0.0	0.0	0.3
All offences	6.4	8.4	6.9	4.8	2.8	1.5	0.6	2.2	5.4
All incidents	3.6	4.6	4.0	3.0	1.8	0.8	0.4	2.2	3.1
Exposure (driver-years×10^3)	49	65	58	74	46	28	14	1	335

Table 5 Percentage of offences committed by unlicensed drivers, 1988-97

Type of offence	*Percent of these offences committed by unlicensed drivers*	*Offences committed by unlicensed drivers as a percentage of all offences*
Insurance Offences	24.5	3.30
Driving while Disqualified	100.0	2.20
Licence Offences[1]	17.5	1.19
Drink or Drugs	13.0	0.94
Speed Limits	1.1	0.51
Theft or Unauthorised Taking	49.9	0.49
Reckless or Dangerous Driving	26.0	0.12
All types		9.30

[1] *includes Provisional Licence offences*

Figure 3 explores the extent to which drivers were disqualified during 1988-97. Section 2.3 explained how the exposure of drivers in terms of the number of driver-years is calculated. The calculation has been repeated counting only periods of disqualification, and the figure presents the number of disqualified driver-years in 1988-97 as a percentage of the number of licensed driver-years. There is a clear peak for male drivers aged 20-24, followed by a steady decline; the pattern is the same for female drivers, but at a much lower level.

3.2.2 Recidivism

One measure of recidivism is the number of offences committed by individual drivers over a specific period, taken here to be the most recent five years. The results in Table 6 relate to drivers with current licences; some will not have been driving throughout the five years, so

Table 6 Percentage of drivers who committed N offences in the past 5 years

	Male drivers		*Female drivers*	
N	*Any*	*Experienced*	*Any*	*Experienced*
0	83.61	82.80	95.02	93.94
1	11.50	12.62	4.27	5.31
2	3.04	3.07	0.54	0.62
3	0.99	0.92	0.11	0.10
4	0.35	0.30	0.03	0.02
5	0.17	0.12	0.01	0.01
6	0.09	0.05	0.01	0.00
7	0.07	0.04	0.00	0.00
8	0.04	0.01	0.00	0.00
9	0.03	0.01	0.00	0.00
>9	0.12	0.04	0.01	0.00
No. of drivers	242723	195744	183792	129074
Mean no. of offences	0.259	0.246	0.060	0.070

'experienced' drivers are defined here to be those who have had a full licence for at least five years. In practice, differences between 'experienced' and 'any' drivers are slight except in the tail of the distribution.

3.2.3 Day of week and week of year

The day of week when an offence was committed can be calculated from the date, also the week of the year. This section compares the distribution of offences in 1993-97 by the day of week and week of year when they were committed. The results are only meaningful for offences which are committed relatively frequently. They are presented as relative rates, i.e. the number committed per unit of time divided by the mean number.

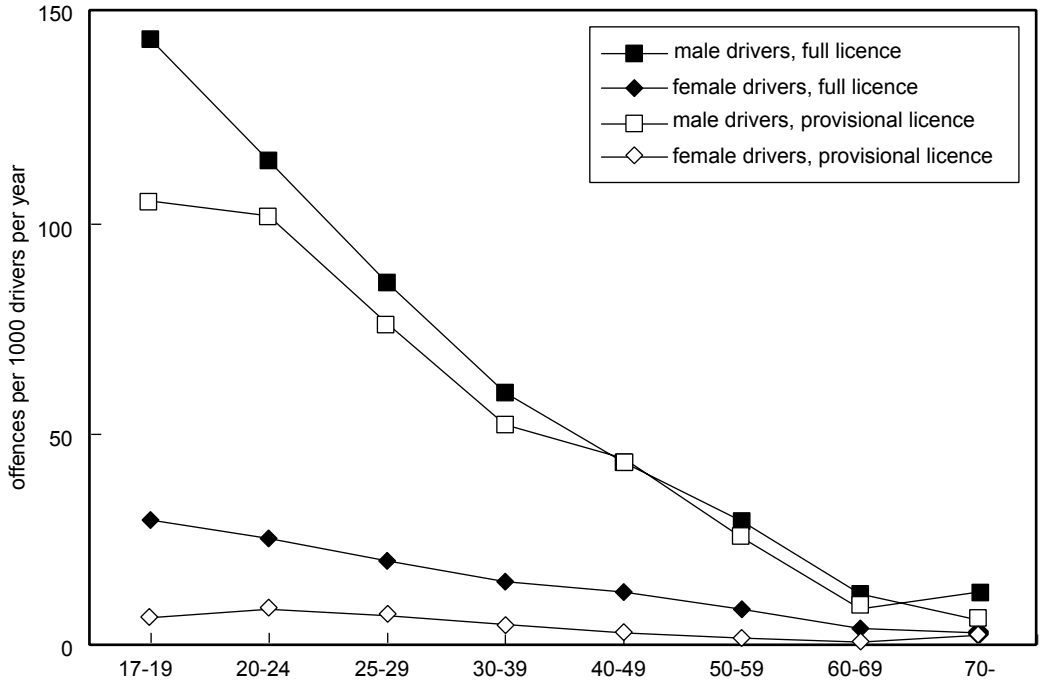


Figure 2 Average number of driving offences per 1000 drivers per year, 1988–97

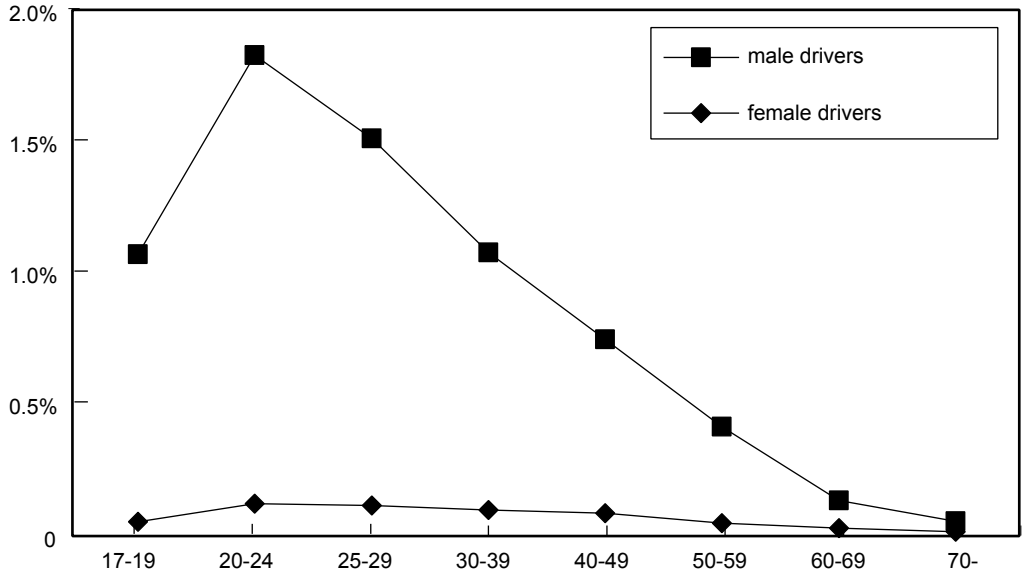


Figure 3 Rate of disqualified drivers, 1988–97

Figure 4 shows the relative rate by day of week for all offences and speeding offences committed by men and women. The relative rates shown for careless driving and drink/driving offences are for offences committed by either men or women, since women commit relatively few of these offences. The relative rate would be 1.0 for each day if equal numbers of offences were committed on each day of the week. Thus, over 40 per cent more drink/drive offences are committed on Sundays than on the average day, and over 30 per cent fewer on Mondays (presumably many of the offences actually committed on a Sunday resulted from drinking on a Saturday evening).

There are clear differences between the six groups of offence. The overall offence rate for men rises steadily from a minimum on Monday to a peak on Saturday, while the overall rate for women falls from a peak on Tuesday to a minimum on Sunday. The most pronounced weekly variation is for drink/driving offences, the rate on Saturday and Sunday being more than twice the rate on Monday–Thursday.

To analyse the distribution of offences by week of year, week 1 is defined as January 1-7, week 2 as January 8-14 etc. December 31 is excluded, also December 30 in the leap year 1996. The weekly totals are rather variable, so totals are then calculated for consecutive 4-week ‘months’.

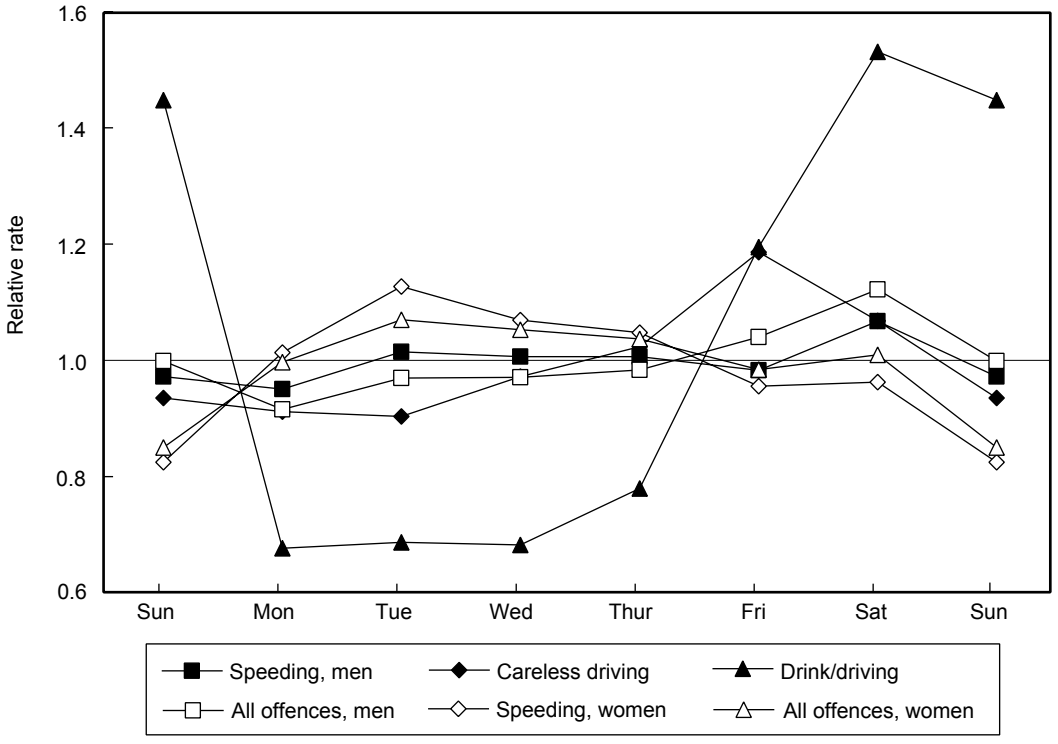
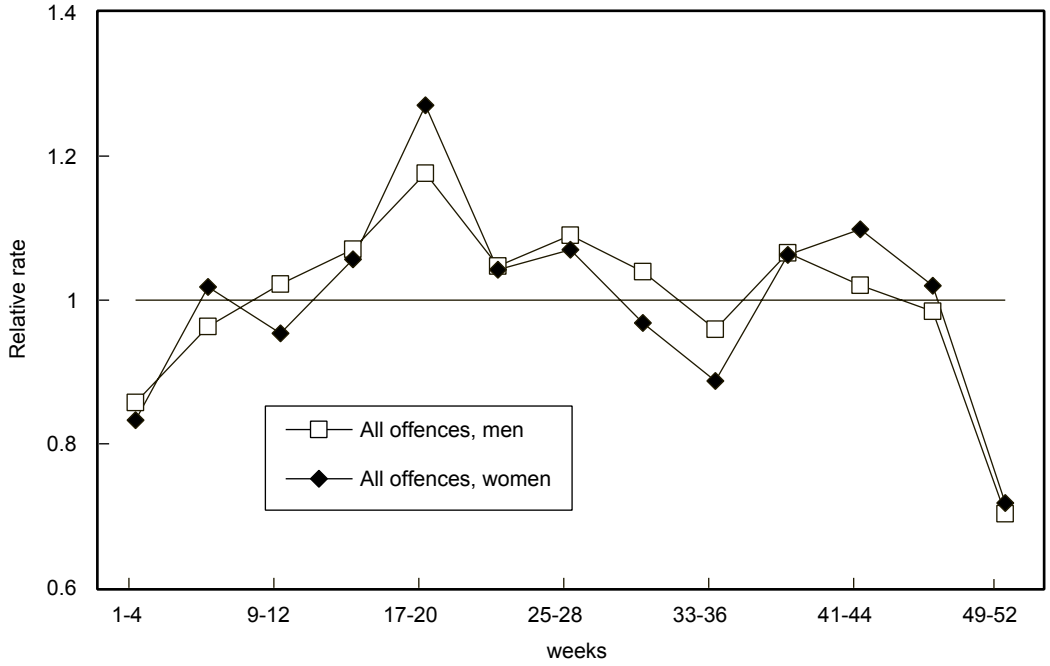


Figure 4 Relative offence rates, by day of week

Figure 5 shows the relative rate per month, so the relative rate for any particular offence type would be 1.0 each month if equal numbers of that offence were committed per month.

The upper set of graphs compares the overall rates for men and women, and shows considerable variation: the graphs for men and women differ. Rates are low in the first four weeks of the New Year and are highest in weeks 17-20 (late April/early May), with a secondary peak for men in weeks 37-40 (September) and for women in weeks 41-44 (October). Rates are lowest in weeks 49-52 (December). The graphs for speeding offences are not shown because they match the upper pair of graphs rather closely, as speeding offences form such a high proportion of all offences.

The lower set of graphs compares the careless driving and drink/driving rates (men and women combined) and these graphs vary less. Drink/driving rises slowly in the early weeks of the year to a minor peak in weeks 21-24 (late May/Early June), with a second peak in weeks 45-48 (November), although even in that 4-week period the rate is only 11 per cent above average.



Note: The graphs for speeding offences closely match these graphs

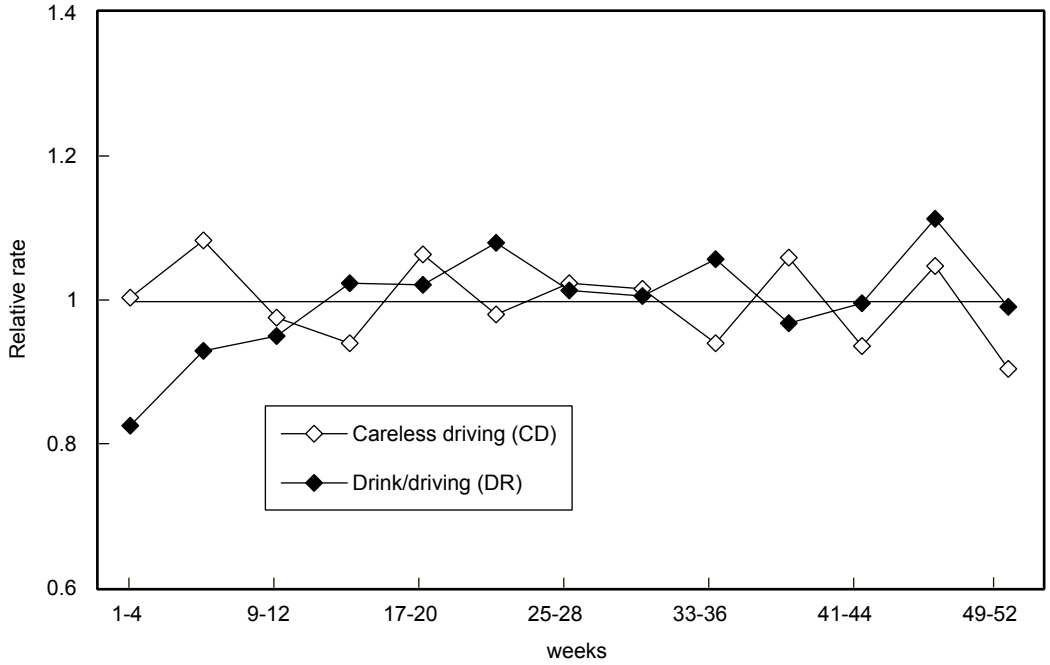


Figure 5 Relative offence rates, by week of year

3.3 Trends in offence rates

The reasons for adopting an indirect approach were described at the beginning of Section 3. This Section describes a simple statistical model that has been developed to examine trends in offence rates.

The model uses the data that were analysed in Tables 2-4, namely the annual number of offences and exposure by driver age, sex and licence-type. It assumes that the offence rate varies from year to year in the same way for each age group for any particular combination of sex and licence-type, although the pattern may well differ between combinations. This assumption leads to a multiplicative model which can be fitted independently for each combination of sex and licence-type:

$$O_{sl}(a,y)/L_{sl}(a,y) = A_{sl}(a) \,.\, Y_{sl}(y) \qquad (1)$$

where $O_{sl}(a,y)$ = number of offences of a particular type committed in year y by drivers of age a, sex s and licence-type l,
$L_{sl}(a,y)$ = number of driver-years of exposure in year y
$A_{sl}(a)$ = factor varying with age a
$Y_{sl}(y)$ = factor varying with year y

The 'age' factor A_{sl} and 'year' factor Y_{sl} are estimated by fitting the model using the GLIM program (Francis et al, 1993). There is no need to compensate for differential rates of name-change, as was necessary for the direct analyses of offence rates, and any changes over time in the composition of the driving population by age and sex are taken into account automatically.

The year factor represents the trend in offence rates over time, the age factor is equivalent to the information shown in Figure 2 but with allowance made for any changes over time. The fit of the model to a particular data set will show whether the multiplicative assumption of model (1) is acceptable in that case.

The standard errors of the estimated coefficients will indicate the likelihood that any apparent difference between years or age-groups might have occurred by chance. For each of the data sets to which the model was fitted, the standard error of the year factor $Y_{sl}(y)$ varies very little with y, so a single 95% confidence interval can be calculated for all years: this shows the interval centred upon the estimated factor within which it is 95% certain that the actual value lies. By contrast, the standard error of the age factor $A_{sl}(a)$ does vary strongly with a, so these intervals cannot be presented as simply. As it is the trend through time that is of particular interest, confidence intervals will be only presented for the year factor.

3.3.1 Trends for all offences

The model is fitted first with O='All offences', and the fit of the model will be considered before the age and year factors. Table 7 shows the 'adjusted R^2' values, i.e. the R^2 values adjusted to allow for the number of variables used when the model was fitted (which is the same for each of the four data sets). It also includes the number of offences for each group of drivers, choosing 1996 as an example, to investigate variations among the adjusted R^2 values. The model fits very well where the number of offences is greatest, men with full licences, and the fit declines as the number of offences reduces. This suggests that the multiplicative assumption is appropriate, and that the relatively poor fit among women with provisional licences is the result of variability caused by the small number of offences committed by this sample of drivers.

Table 7 Fit of model (1), all offences

	Adjusted R^2 values		*Number of offences in 1996*	
	Men	*Women*	*Men*	*Women*
Full licence	0.994	0.964	11082	2480
Provisional licence	0.929	0.531	1055	124

The age and year factors estimated by the model are shown in Figures 6 and 7. A specific year or age-range must be selected to display the results. The year factor for the 20-24 age group is displayed in Figure 6 and the age factors for 1996 are displayed in Figure 7, but the same patterns would be found in the graphs based on any other choices.

Figure 6 shows an increasing trend among men with full licences until 1989, and subsequent fluctuations. The 95% confidence interval for this graph is ±4.3, so these changes are measured rather precisely. The 95% confidence interval for the trend for men with provisional licences is ±14.8 so that, although the offence rate has clearly fallen rather steadily, some of the detailed changes shown may have occurred by chance. Among women with full licences, the 95% confidence interval is ±2.0, so it is clear that the trend almost doubled between 1985 and 1997. The 95% confidence interval is ±3.6 among women with provisional licences, so the fluctuations of this graph may well have occurred by chance; in effect, the trend has been static.

Figure 7 generally resembles Figure 2, except for the relationship between the graphs for men with full and provisional licences. The reason is that Figure 7 presents the situation in 1996 while Figure 2 presents the 1988-97 average; Figure 6 shows that the rate for the latter group had fallen well below its 1988-97 average by 1996, while the rate for the former group remained about its 1988-97 average.

3.3.2 Male drivers with provisional licences

The most striking feature of Figure 6 is the steady reduction since 1988 of the offence rate of male drivers with provisional licences. This is examined in more detail in Figure 8, which shows the actual annual offence rates for the 17-19 year old subset of these drivers (the age group with the greatest exposure in terms of driver-years).

Only one technical change is known to affect these rates: most UT offence types ceased to be endorsable from 30 June 1992, so the reduction from about 11 UT offences per thousand drivers per year before 1992 to 2 from 1993 does not reflect a genuine reduction in this type of offence. The replacement of Provisional Licence (PL) codes by LC

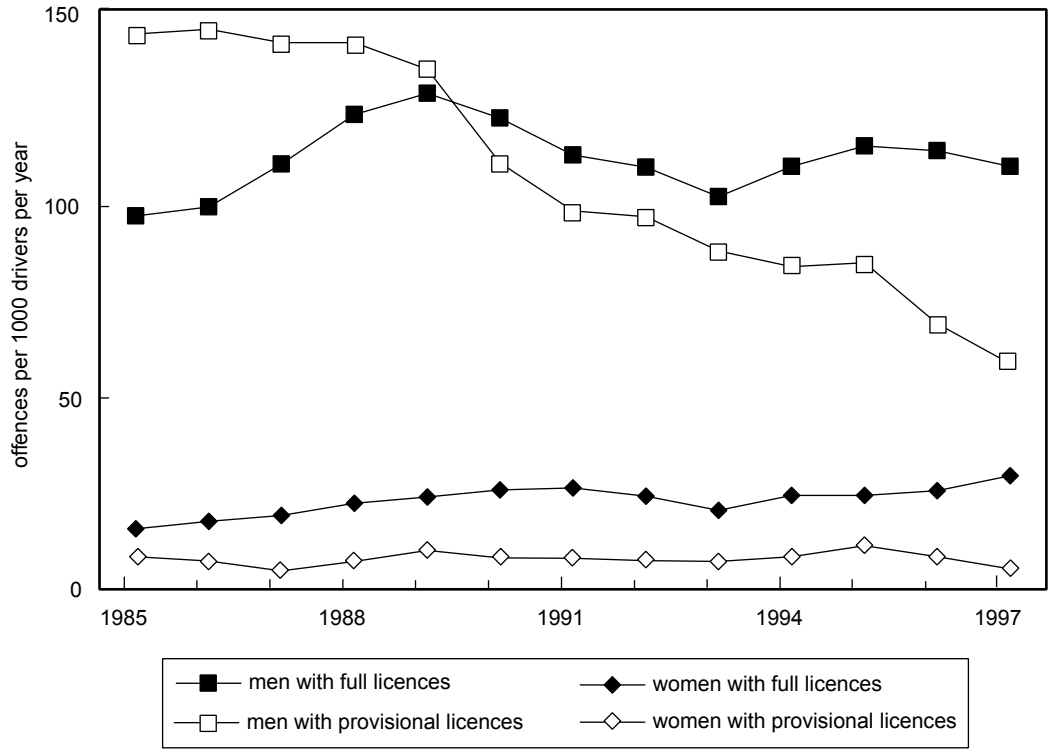


Figure 6 Annual number of driving offences per 1000 drivers, 1985–97 (Results from model (1), standardised for 20–24 year olds)

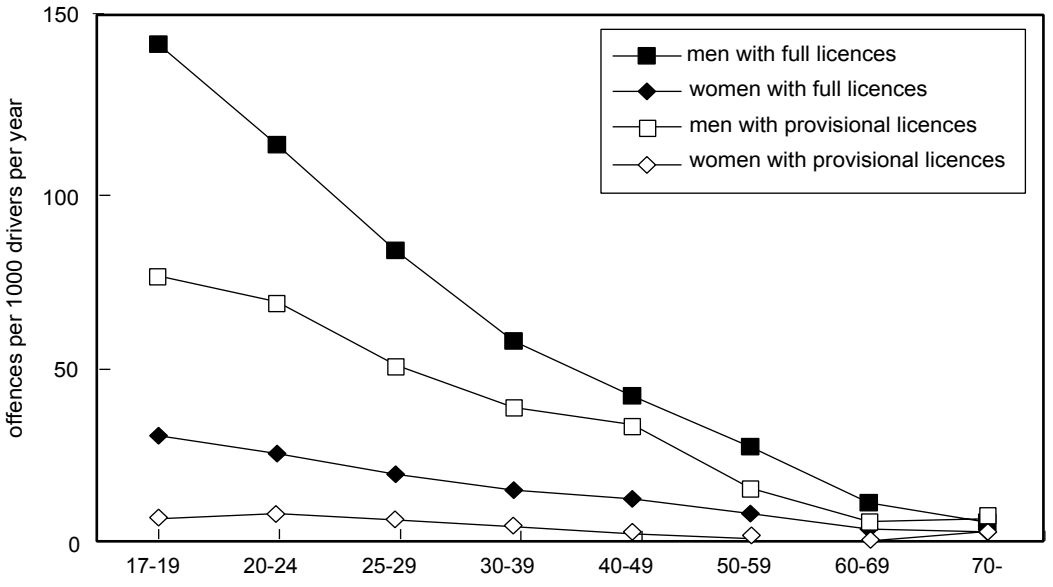


Figure 7 Annual number of driving offences per 1000 drivers, by age (Results from model (1), standardised for 1996)

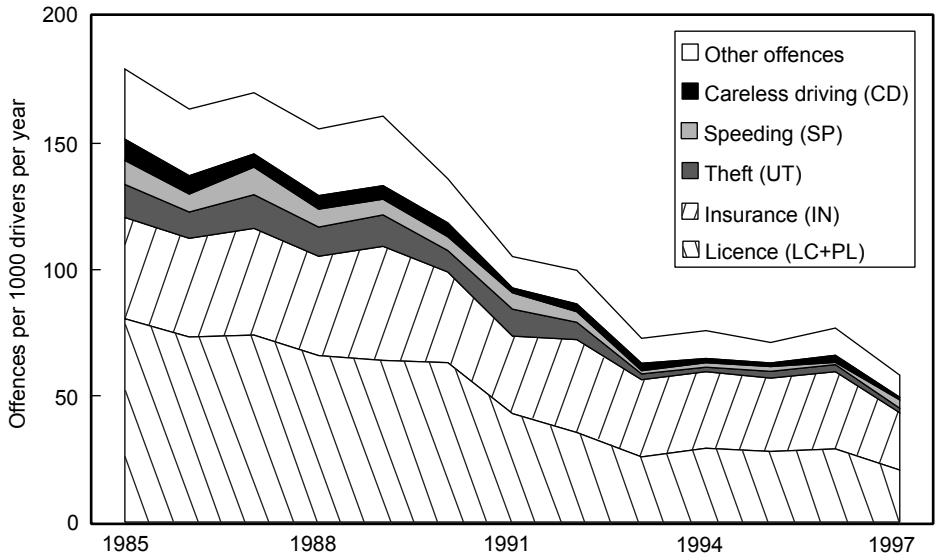


Figure 8 Annual number of driving offences per 1000 male drivers with provisional licences aged 17–19

codes in 1992 is accommodated by adding PC and LC offences together throughout.

Apart from the change to the UT offences, the reductions in the major groups of offences are probably the result of a reduction in offending behaviour. The main contribution to the reduction in the overall rate has come from the rapid fall in the rate of Licensing offences, which has been less than the rate of Insurance offences each year from 1992. A similar pattern is found among the older age groups.

3.3.3 Trends for three types of driving offence

Several of the DVLA offence types appear to be particularly relevant to road safety, but offences such as Dangerous Driving are committed rather infrequently so trends cannot be measured reliably with a small sample of licence records. Three offence types can be studied using model (1): Speeding, Careless Driving and Drink/Driving. Figure 9 brings together the year and age factors for these three, and also shows the 95% confidence intervals for the year factors.

The trends for Speeding offences among men and women with full licences have risen steadily for much of the period studied. Offence rates are much lower among drivers with provisional licences, indeed the trend has fallen for men and is effectively zero for women. The rates fall steadily with age among drivers with full licences, but are generally much lower among those with provisional licences and vary little with age.

The trends for Careless Driving offences have generally fallen, except for a peak for men with full licences in 1988. The rates fall rapidly with age among those with full licences; again, they are much lower among drivers with provisional licences and vary little with age.

The trend for Drink/Driving offences among men with full licences generally fell between 1989 and 1995, but has been slightly higher in 1996-97; among men with provisional licences, however, the trend has fallen reasonably steadily throughout. Offence rates have remained much lower among women. The rates for men with full and provisional licences rise at first with age, and converge in the 30-39 age group.

Thus, these three types of offence have followed three different types of trend, with three different patterns of variation by age.

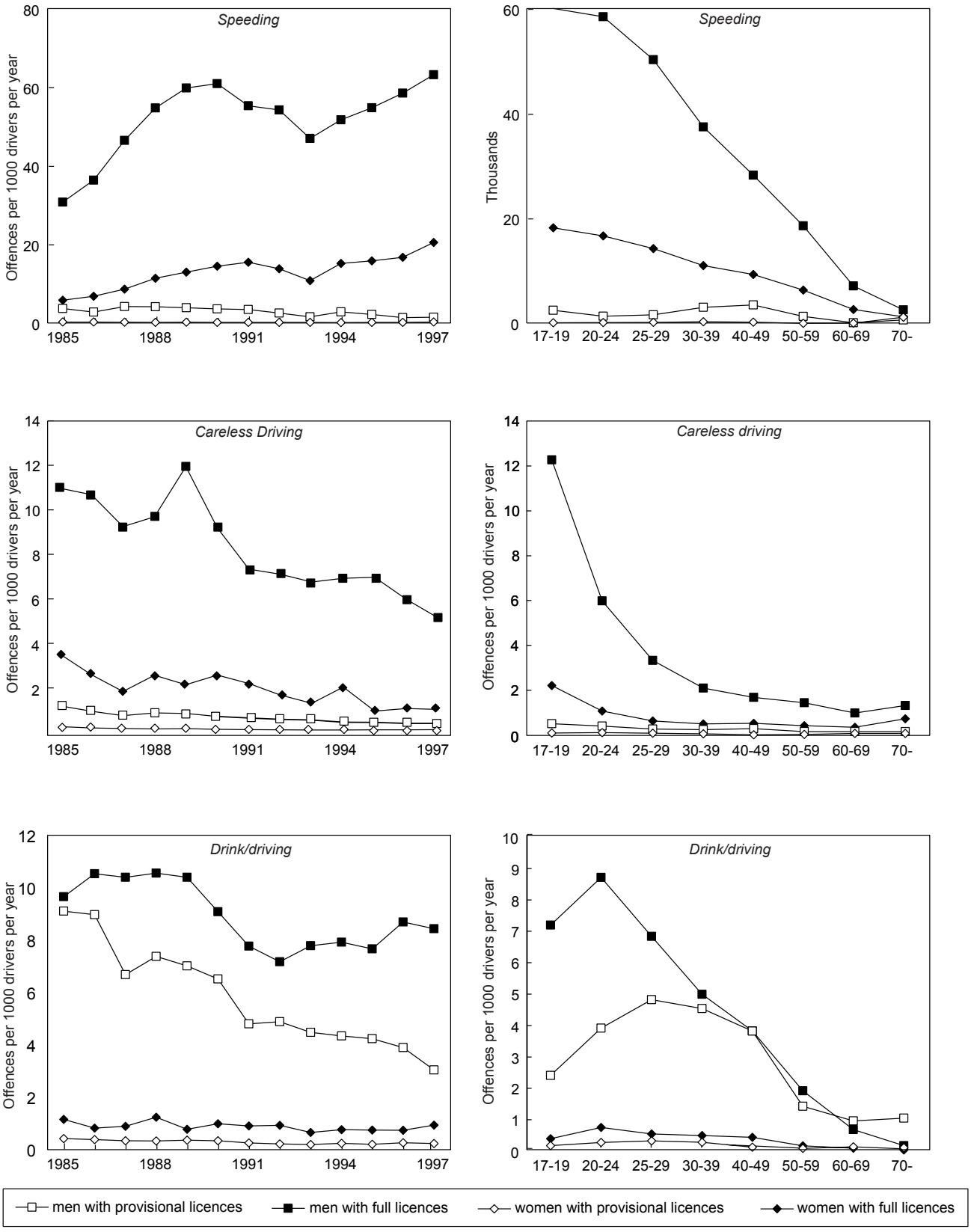
3.4 Regional comparisons of offence rates

Section 2.1 explained that one of the details in the licence record is the driver's postcode. It is sometimes out of date because drivers often fail to advise DVLA of a change of address, but the region of residence will still be correct if the driver had moved house within the region. Thus, the DVLA records should show the region with sufficient accuracy to allow regional differences to be analysed. Eleven regions are used: Wales, Scotland and the current Government Office Regions in England (except that Merseyside has been combined with North West England). The regions are:

Scotland	North East England
North West England	Yorkshire & Humberside
East Midlands	West Midlands
Wales	London
South East England	South West England

There are clear differences in the regional distributions of drivers by age and sex, and these must be taken into account to obtain a fair comparison. For example, the population of South West England is relatively elderly, which tends to give a favourable impression of the overall offence rate in that region because it has been seen that offence rates generally fall with age of driver. The basis for comparing regions will be the offence rate per thousand driver-years by age and sex, so differences in regional population will be taken into account automatically.

The model of regional rates is similar to (1) and will be applied to all licensed drivers in each region: full and provisional licence-holders are grouped together because the numbers of drivers in some regions are relatively small. Also, to obtain a general impression of whether rates are rising or falling in a particular region, rates in



95% confidence intervals for graphs by year

	Full licences		*Provisional licences*	
	Men	*Women*	*Men*	*Women*
Speeding	±3.4	±1.3	±1.3	±0.2
Careless driving	±0.6	±0.4	±1.8	±0.2
Drink/Driving	±0.8	±0.3	±1.6	±0.2

Figure 9 Annual number of driving offences per 1000 drivers, for 3 offence types (Results from model (1), standardised for 20–24 year olds and 1996)

1993-97 will be compared with rates in 1988-92. The regional model is fitted for men and women separately, and can be expressed:

$$O_s(a,r,p)/L_s(a,r,p) = A_s(a) \,.\, R_s(r,p) \qquad (2)$$

where $O_s(a,r,p)$ = number of offences of a particular type committed in period p (1988-92 or 1993-97) by drivers of age a and sex s living in region r,
$L_s(a,r,p)$ = number of driver-years of exposure
$A_s(a)$ = factor varying with age a
$R_s(r,p)$ = factor varying with region r and period p

The main interest lies in $R_s(r,1993\text{-}97)$, the offence rate per region in the last five years. The relation between this and $R_s(r,1987\text{-}92)$ indicates whether rates have tended to fall or rise between the two periods: the possibility of using finer time bands will be examined later. The main role of A_s is to allow for any differences between regional age distributions.

Model (2) is fitted first with O='All offences', and its fit will be examined to see whether the multiplicative assumption is acceptable. The model has more variables to fit than the trend model (1) and the number of offences for each combination of a, r and p tends to be smaller, so the adjusted R^2 values would be expected to be lower. Nonetheless, values of 0.974 for men and 0.857 for women compare favourably with Table 7, indicating that the assumption appears to be reasonable once again.

Figure 10 shows the regional factors $R_s(r,1993\text{-}97)$ and again the 20-24 year age group has been chosen to present the results. In this figure, three bars are superimposed for each region: the shortest and longest show the 95% confidence intervals while the central bar shows the best estimate of the regional rate. Thus, the best estimate of the offence rate of 20-24 year old male drivers in 1993-97 in North West England was 142 per thousand driver-years; the model shows that one can be 95% certain that the true rate lay between 134 and 151. Comparing adjacent bars, it is highly likely that the rate was actually higher in North West England than in Wales, and virtually certain that it was higher than in any other region of Great Britain.

The confidence intervals are relatively wide for women drivers, probably because of the lower rates. Nonetheless, it is again highly likely that the overall rate was highest in North West England, although in this case the second highest rate appears to have been in South East England.

The analyses show clearly that offence rates vary widely between regions. There are various possible reasons in addition to driver behaviour, including police enforcement and Court-related factors such as the willingness of magistrates to convict.

Figure 11 compares the offence rates in 1993-97 to those in 1988-92, based on the R_s factors for the two periods, so results below the 100% line indicate that rates were lower in 1993-97 than in 1988-92. As with the previous figure, the best estimates are presented together with the 95% confidence intervals. The intervals are relatively wider than in the previous figure, and their width

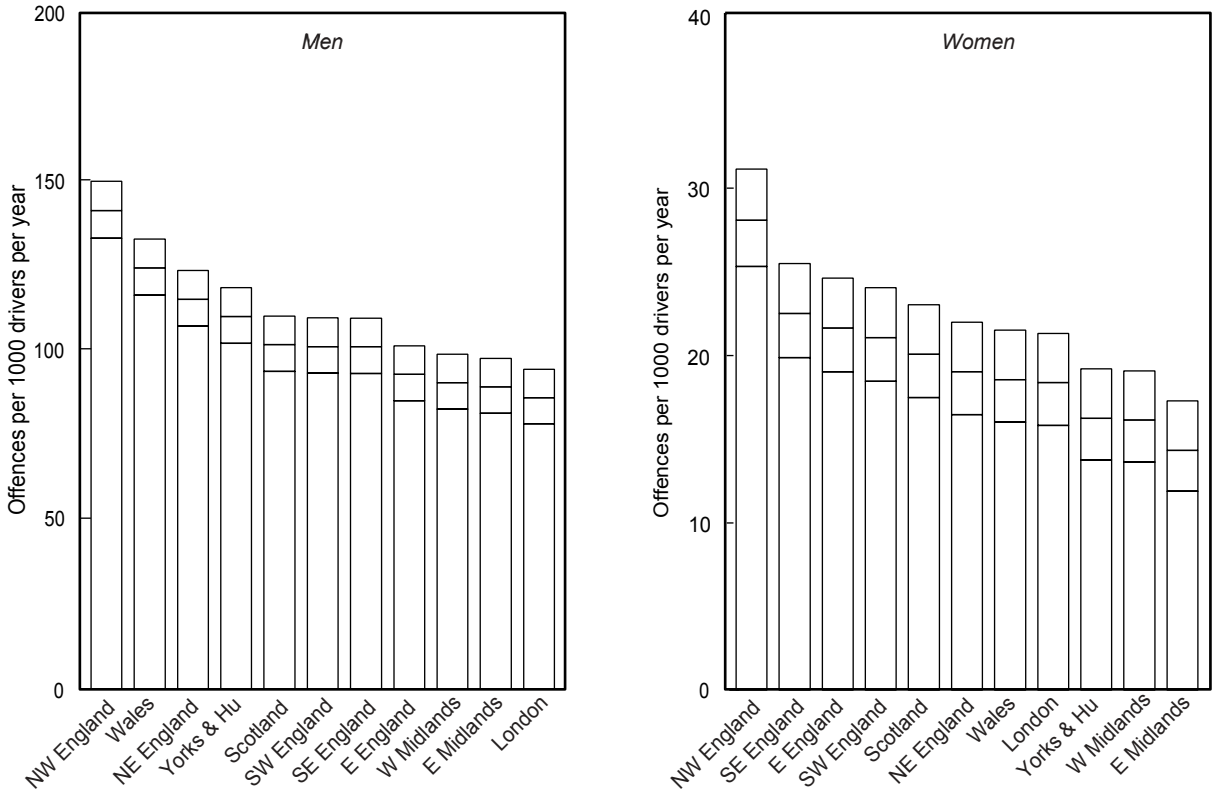


Figure 10 Offence rate per thousand driver–years, 1993–97, by region (Results from model (2) for 20–24 year old drivers, 95% confidence intervals included)

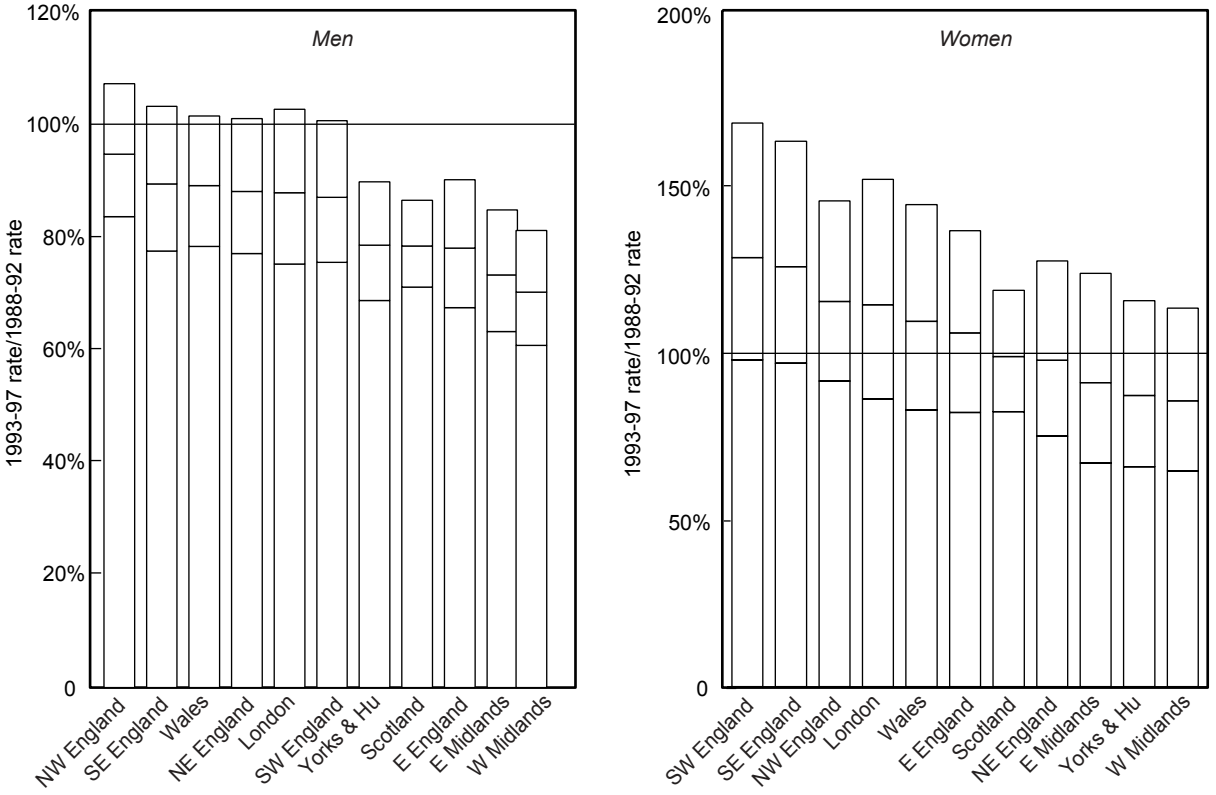


Figure 11 Offence rate in 1993–97/Offence rate in 1988–92, by region (Results from model (2), 95% confidence intervals included)

indicates that no more detailed regional analysis of trends through time could usefully be made since shorter periods would lead to even wider confidence intervals.

The figure shows that rates for men in North West England have scarcely changed, but the rate has probably fallen in other regions and it has almost certainly fallen in about half the regions: the greatest fall appears to have been in the West Midlands, by nearly three-tenths. The confidence intervals for women are wider, but again the rate appears to have fallen most in the West Midlands - although even here it is possible that the rate actually rose. The greatest increase appears to have been in South West England.

It is interesting to compare these regional rankings with those found by Broughton (1986) for 1983. The analytical method and presentation of the results differ slightly from those used previously, but it appears that in 1983 the overall offence rate in North West England was only slightly above the national mean, and rather less than in North East England (referred to as Northern England in the earlier report) which had the highest rate. This is consistent with the results of Figure 11 which show that male rates have fallen less over the past decade in the North West than in other regions such as the North East. The relatively high overall rate in North West England appears to be a fairly recent phenomenon.

Figure 12 examines whether the results for men and women in the previous two figures are correlated, since it appears likely that specific regional influences would affect male and female rates similarly. The comparison is complicated by the confidence intervals: the graphs show points representing the best estimates for men and women, but ideally would show ellipses representing joint confidence intervals. The left-hand graph compares the 1993-97 regional offence rates, and appears to show a weak positive correlation; however, this is almost entirely due to the results from North West England and r^2 falls from 0.36 to 0.02 when that point is removed. The right-hand graph compares the relative rates by region, and the evidence is much clearer here: r^2=0.58 with no clear outliers.

Thus, it appears that the factors which lead to more or fewer driving offences in one region than another during one period do not affect men and women equally, but they do lead to similar changes for men and women over time.

To examine possible reasons for the lack of a correlation between the regional offence rates for men and women in 1993-97, Figure 13 compares the regional rates (number per thousand drivers of all ages per year) for the more common groups of offence among men and women; it shows considerable variation between men and women. Since different factors are likely to influence the rate at which different types of offence are committed, the differences between the two sets of distributions may partially account for the lack of correlation. A more sophisticated analysis of regional variations in some offence rates is presented in the next section.

3.4.1 Comparisons for three types of driving offence

In principle, model (2) can be applied with any group of driving offences. In practice, however, as the number of offences in the group examined falls, the confidence intervals grow wider relative the offence rate and it becomes more difficult to draw reliable conclusions. This is

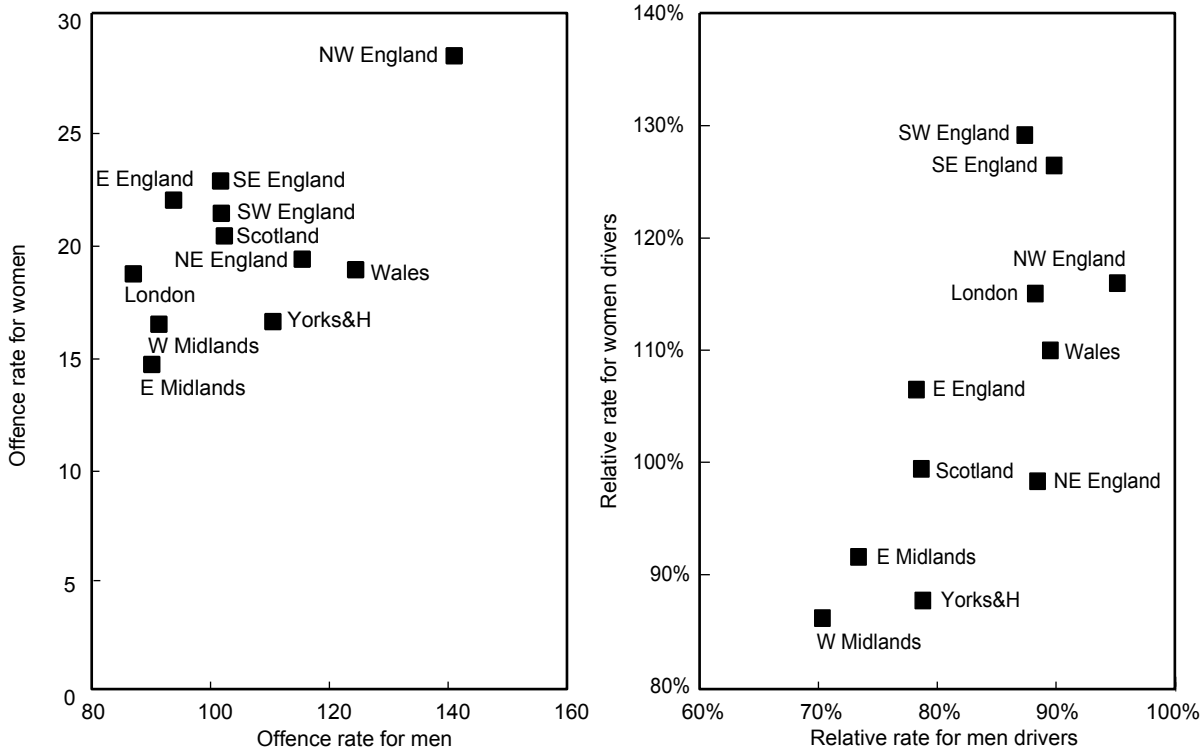


Figure 12 Relatinship between reginal rates for men and women

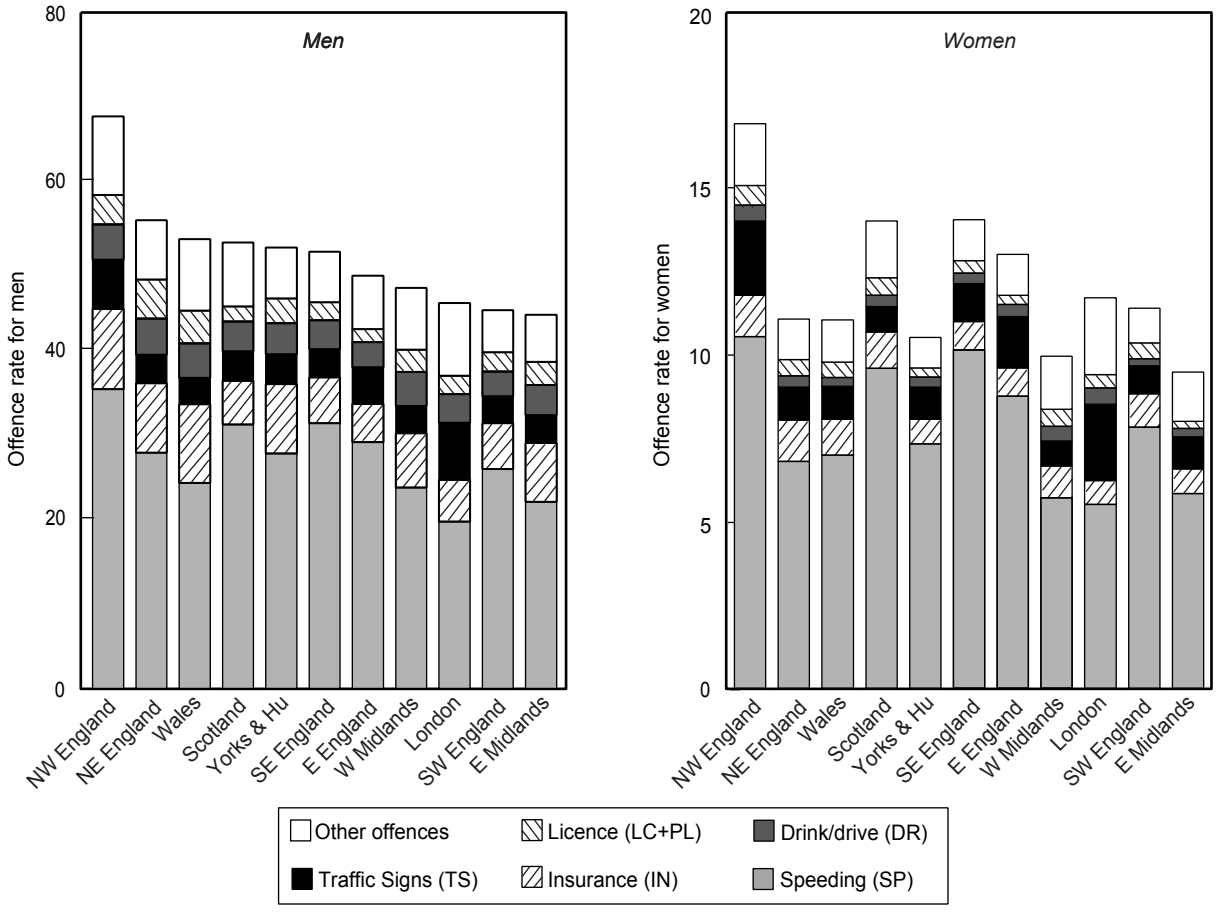


Figure 13 Regional offence rates for men and women, 1993–97, by type

especially true with women drivers, so Table 8 presents results for Speeding offences (men and women) and for Careless driving and Drink/driving offences (men only). In each part of the table, regions are listed by decreasing order of the best estimate of the 1993-97 offence rate. If the ratio of a 1993-97 rate to a 1988-92 rate is shown to be greater than 100%, the rate has tended to grow over the past decade, whereas if it is less then the rate has tended to fall.

The Table illustrates the increasing difficulty of making reliable comparisons as the incidence of offences decreases: for example, the 95% confidence interval for the drink/driving offence rate in the East Midlands includes nine regions. Nevertheless, some reasonably clear regional differences do emerge, with the highest regional rates of speeding offences being almost twice the lowest.

The male and female speeding rates by region correlate rather well (r^2=0.79 with NW England, 0.70 without), in contrast to the lack of correlation found previously for All offences. This supports the idea that the lack of correlation is caused at least partly by varying regional distributions of offence types.

4 The number of licensed drivers

The 'exposure' measure used in Section 3 to calculate offence rates for the drivers in the TRL sample has an independent value, for it indicates the development of the driving population over the years covered by the TRL data. The main problem with this source arises from the loss of information for earlier years as a result of name-changes, as mentioned in Section 2.2. The exposure series for women can be adjusted to compensate, as described in the Appendix; the loss is much less significant for men, so no adjustment has been made. The adjustment cannot be exact, but should allow the exposure data from earlier years to be reconstructed sufficiently well for the general trends to be established.

The number of driver-years, $L_{sl}(a,y)$ from Section 3.3, is analysed first to study trends in numbers of licensed drivers. Figure 14 shows how the number of drivers with full licences grew between 1985 and 1997: by 22 per cent overall for men and 53 per cent for women. Among men, the growth was very similar for provisional licences, 23 per cent, but much less among women - 33 per cent.

Figure 15 repeats Figure 14 for six age ranges. It shows strong growth rates among the 40- age group, less strong growth in the 30-39 age group and some recent declines in the younger age groups. The latter are probably the result of demographic changes caused by fluctuating birth rates in the 1960s and 1970s, rather than any diminution in the desire of young people to drive.

The consequence of these changes is illustrated by Figure 16 for drivers with full licences. Within the total of these drivers, the proportion of older drivers and women has grown so that in 1996 the number of women at least 50 years old exceeded the number of men aged 17-29.

Table 8 Regional rates for speeding, careless driving and drink/driving offences

	Speeding offences, men			*Speeding offences, women*	
	1993-97 rate	*1993-97 rate / 1988-92 rate (%)*		*1993-97 rate*	*1993-97 rate / 1988-92 rate (%)*
NW Eng	58 (55,62)	108 (96,121)	NW Eng	15 (14,17)	128 (105,157)
SE Eng	51 (48,55)	120 (106,136)	SE Eng	14 (13,16)	172 (137,217)
E Eng	48 (45,51)	98 (86,110)	Scotland	13 (11,14)	101 (88,117)
Scotland	47 (44,50)	78 (72,85)	E Eng	12 (11,14)	120 (97,150)
Yorks & H	46 (43,49)	85 (75,97)	SW Eng	12 (11,13)	139 (110,176)
NE Eng	46 (42,49)	109 (95,124)	Wales	11 (9,12)	218 (158,301)
SW Eng	44 (41,47)	98 (86,112)	Yorks & H	10 (9,12)	107 (85,135)
Wales	39 (36,42)	89 (78,102)	NE Eng	9 (8,11)	123 (94,160)
W Mid	38 (35,42)	91 (79,104)	London	8 (7,9)	132 (97,179)
E Mid	35 (32,39)	82 (71,94)	W Mid	8 (7,9)	90 (69,117)
London	31 (28,35)	99 (84,117)	E Mid	8 (7,9)	94 (72,123)
	Careless driving offences, men			*Drink/driving offences, men*	
	1993-97 rate	*1993-97 rate / 1988-92 rate (%)*		*1993-97 rate*	*1993-97 rate / 1988-92 rate (%)*
NW Eng	7.5 (6.8,8.3)	72 (60,86)	NE Eng	8.5 (7.6,9.6)	81 (64,103)
Wales	6.5 (5.8,7.3)	72 (59,87)	NW Eng	8.3 (7.4,9.4)	87 (68,111)
NE Eng	6.0 (5.3,6.8)	87 (71,108)	Wales	8.3 (7.3,9.3)	100 (78,128)
W Mid	5.8 (5.1,6.6)	62 (51,76)	W Mid	8.2 (7.3,9.3)	79 (62,100)
Scotland	5.7 (5.0,6.5)	51 (44,58)	SE Eng	7.7 (6.7,8.7)	98 (76,126)
London	5.6 (4.9,6.4)	72 (59,89)	Yorks & H	7.6 (6.7,8.6)	73 (57,93)
E Eng	5.3 (4.6,6.1)	69 (56,86)	E Mid	7.5 (6.6,8.5)	88 (69,114)
SW Eng	4.5 (3.9,5.3)	82 (64,105)	SW Eng	6.7 (5.8,7.8)	84 (65,110)
Yorks & H	4.5 (3.8,5.3)	62 (49,79)	Scotland	6.6 (5.7,7.6)	82 (68,99)
E Mid	4.4 (3.8,5.2)	63 (50,80)	E Eng	6.6 (5.7,7.6)	82 (63,106)
SE Eng	4.1 (3.5,4.9)	55 (43,70)	London	5.8 (4.9,6.8)	78 (59,103)

1993-97 rates are for 20-24 year olds results are best estimates, 95% confidence intervals are shown in brackets

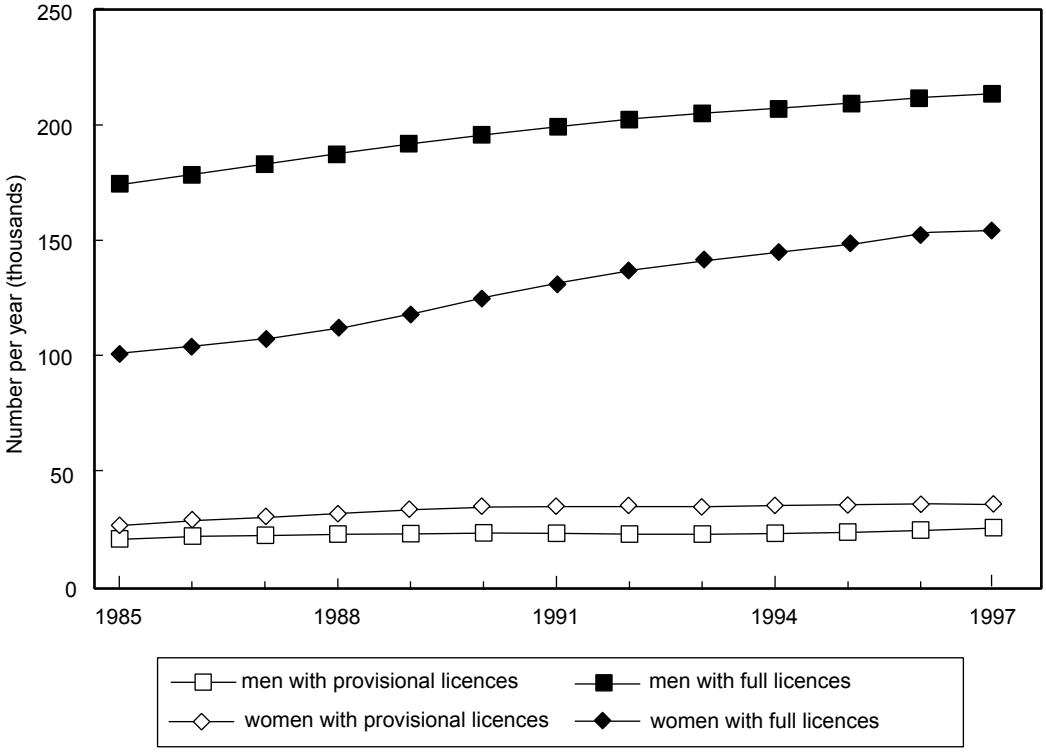


Figure 14 Number of licensed drivers in sample, 1985–97

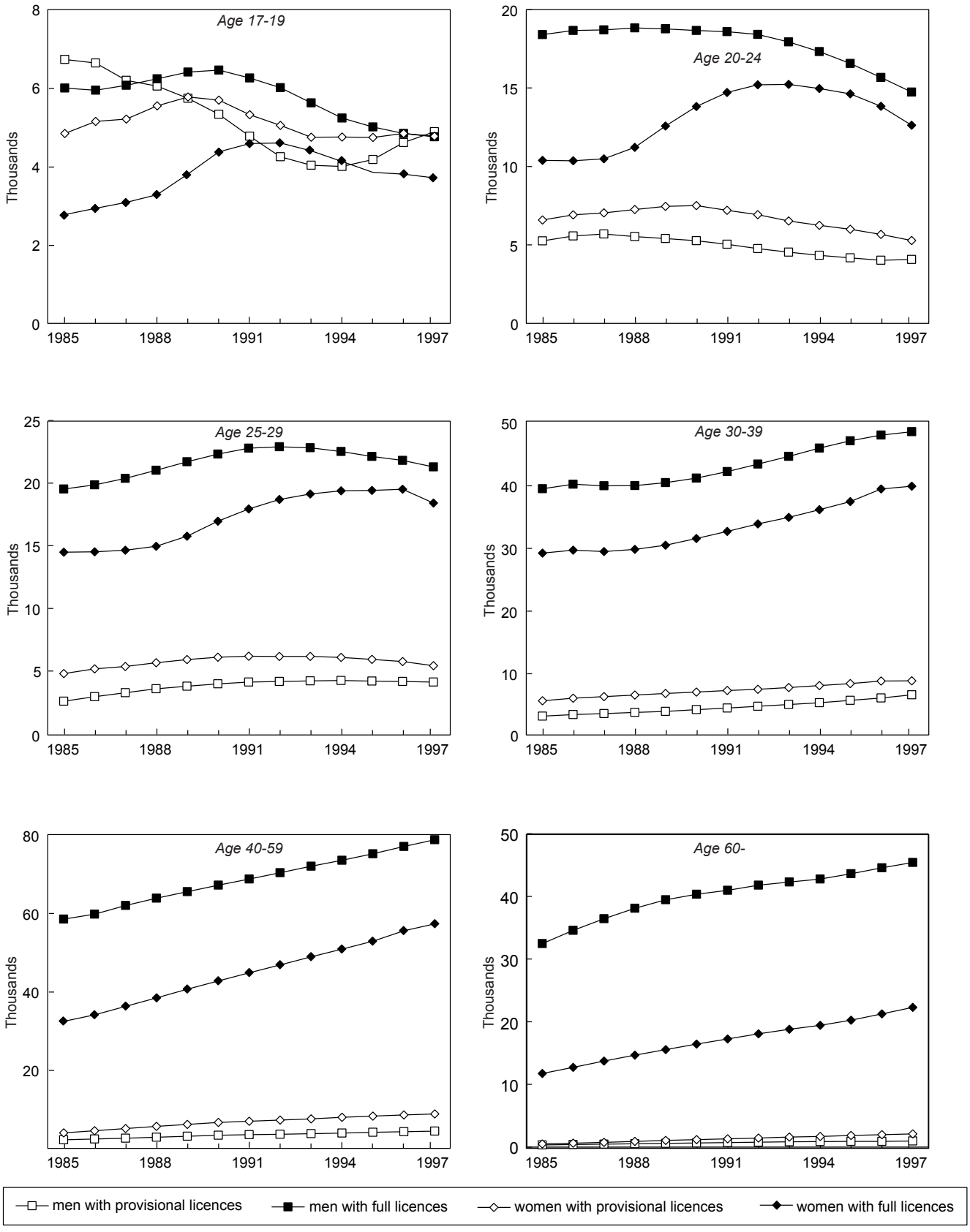


Figure 15 Number of licensed drivers in sample per year, 1985–97, by age

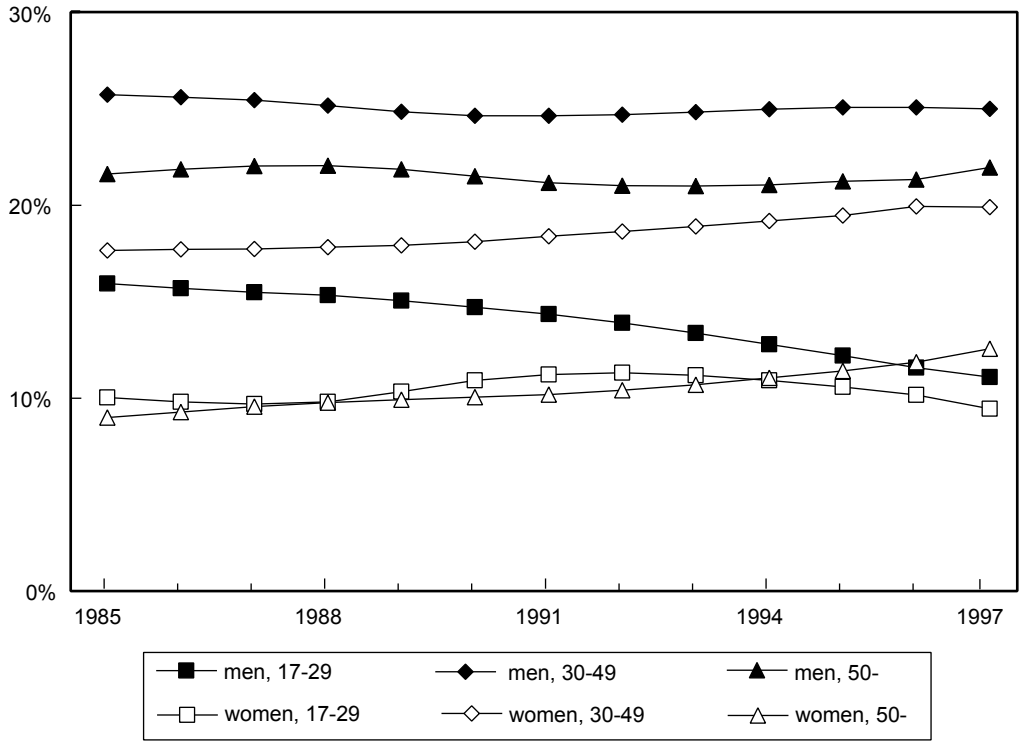


Figure 16 Distribution of drivers with full licences, by age and sex

Possible regional variations in regional licensing trends can be examined with the regional exposure data used in Section 3.4 to calculate regional offence rates. There are few clear differences, most could have been the effect of chance on the sample of drivers. Among men, growth between 1988-92 and 1993-98 appears to have been greatest in London, followed by Scotland, and least in South West England. Among women, growth appears to have been greatest in Scotland, and again least in South West England.

5 Conclusions

This report has analysed the distribution of driving offences among the driving population of Great Britain, using data from the DVLA's Driver Licence file. Many details are routinely removed from the DVLA file, in accordance with the relevant legislation, which means that the DVLA records of many drivers are incomplete. The removal of this information causes problems when trying to analyse trends in offence rates, and an archive of licence information has been set up at TRL to overcome these: it contains details of approximately 1 per cent of licensed drivers which are essentially complete since 1984.

Comparisons of the number of licences recorded in the TRL archive with the numbers recorded in the full DVLA file at various dates in the past shows that the TRL sample of male drivers has grown in line with the full population of drivers. This is not true for female drivers, however, many of whom change their surnames following marriage and consequently leave the sample of drivers represented in the archive. It has been possible, nevertheless, to design analyses which compensate for this loss, so that the patterns which have been identified should not be influenced by the steady loss of female drivers from the archive.

These analyses focus upon the number of offences per thousand driver-years by age and sex of driver over the period 1985-97. The DVLA file only contains details of those driving offences which are sufficiently serious to require the offender's driving licence to be endorsed, so lesser offences such as parking and obstruction have not featured in the analyses.

The rate of driving offences has been seen to vary considerably by age and sex of driver. Men commit many more offences than women, and rates tend to fall with age. The overall offence rate rose in the late 1980s among men with full driving licences, but has since changed little. The overall rate among men with provisional licences, however, has fallen steadily in the 1990s, and is now well below the rate among men with full licences. The chief factor contributing to this fall has been the rapid reduction in the rate of Licensing offences.

Three groups of offences that are particularly relevant to road safety have also been studied: Speeding, Careless Driving and Drink/Driving. Different patterns were found in each case. The rate of Speeding offences per thousand drivers with full licences rose through most of the period studied, but fell among drivers with provisional licences; in any particular year, the rate fell gradually with age. The rate of Careless Driving offences fell through most of the period studied, and in any single year fell rapidly with age. The rate of Drink/Driving offences among male drivers with full licences fell until 1993 but has since risen slightly; the rate has fallen in most years among male drivers with provisional licences and has remained low

among female drivers. In any single year, the rate of these offences among male drivers with full licences is greatest in the 20-24 age group, but in the 25-29 age group for those with provisional licences.

Offence rates were also examined by region, looking both at average rates for the 1993-97 period and changes since the 1987-92 period. The highest 1993-97 rates for male and female drivers were found in North West England; several regions have rather lower rates, the lowest male rate was found in London and the lowest female rate in the East Midlands. Male rates appear to have fallen least in North West England and most in the West Midlands; female rates appear to have risen most in South West England and fallen most in the West Midlands.

The relatively small number of offences in the TRL archive from some regions limits the comparisons that can usefully be made of regional rates for specific types of offence. It does appear, however, that rate of Speeding offences is highest in North West England, almost twice as high as in London. A similar range was found with Careless Driving offences by men, but in this case the lowest rate was in South East England. The range was much less with Drink/Driving offences.

Many drivers commit no offences, a few commit many offences. Among drivers who have held a full licence for more than five years, 82.8 per cent of men and 93.9 percent of women had committed no offences in the past five years. 0.04 per cent of men had committed at least 10 offences, but effectively no women. However, when all drivers are considered (i.e. drivers with a provisional licence or no licence are also included), the percentage rises to 0.12 for men and 0.01 for women.

6 References

Broughton J (1986). *Analysis of motoring offence details from DVLC driving licence records*. Research Report RR77: Transport Research Laboratory, Crowthorne.

Francis B, Green M and Payne C (eds) (1993). *The GLIM System, Release 4 Manual*. Clarendon Press, Oxford.

Wilkins G and C Addicott (1998). *Motoring offences, England and Wales 1996*. Home Office Statistical Bulletin 8/98: Home Office, London.

7 Acknowledgements

The work described in this report was carried out in the Safety and Environment Resource Centre of TRL.

Appendix A: Eliminating bias due to name-changes

Section 2.3 explained how potential bias of the offence rates due to the loss of drivers from the TRL archive as a result of name-changes can be avoided. This involves analysing groups of driver with fairly narrow age ranges so that the effects apply consistently throughout the group. The following example shows how bias can arise when aggregate rates are calculated: it is based on the pattern found among women drivers, but the values are purely illustrative.

	Actual details		*Percent of records lost*	*Details recorded in file*	
	Drivers	*Offences*		*Drivers*	*Offences*
Young	100	100	50	50	50
Middle-aged	200	100	20	160	80
Older	400	80	0	400	80
Total	700	280		610	210
Rate	280/700=0.40			210/610=0.34	

Rates for combined groups of women drivers need to be adjusted to compensate for the effects of differential loss of records, but it was seen that the loss was so small for men drivers that no adjustment was needed. The adjustment factors are derived as follows:

O(a,y) = recorded number of offences committed by women drivers of age a in year y;
E(a,y) = recorded exposure of these drivers;
L(a,y) = proportional loss of records for these drivers
W(a,y) = 1/(1-L(a,y)).

Since the loss of records has probably affected the offences and exposure equally then

- the actual value of O(a,y) was O(a,y).W(a,y) and
- the actual value of E(a,y) was E(a,y).W(a,y)

Hence, the 'true' rate for some grouping of a and y is

$$\Sigma_{a,y} O(a,y).W(a,y) / \Sigma_{a,y} E(a,y).W(a,y)$$

The same weighting method could be used to adjust individual series, such as the offence data shown in Figure 1.

A1 Estimation of weights

The weights W can be estimated from the Sampling Fractions for women drivers in the various age bands. These show that the loss has been greatest among younger drivers, indicating that the principal cause is change of surname following marriage.

The peak age for marriage for women over the past decade has been 26. Accordingly, a simple model was defined: the proportion of women who marry at age A was modelled as

$$P(A) = \alpha.\beta^{A-26} \quad \text{for } A \geq 26 \qquad (3)$$

where a and b are coefficients to be estimated. By experiment, and to avoid introducing another coefficient, P was extended to lower ages by linking the values to those for higher ages:

$$P(A) = P(39-A/2) \text{for } A<26 \qquad (3')$$

This assumes that the propensity to marry and the distribution of age at marriage has not changed over the period studied. Although perhaps not strictly true, changes have occurred only slowly so that any errors introduced into the adjustment process should be trivial.

From these proportions, the probability of a women of age A changing her name within eight years can be calculated (the same equations apply irrespective of the functional form chosen for P). The probability is the sum of:

i the probability that she changes her name in the first year (at age A);

ii the probability that she does not change her name in the first year (at age A) but does change her name in the second year (at age A+1);

iii the probability that she does not change her name in the first or second year (at ages A or A+1) but does change her name in the third year (at age A+2);

...

viii the probability that she does not change her name in the first seven year (at ages A to A+6) but does change her name in the eighth year (at age A+7).

Writing P_A in place of P(A) for clarity, this is equal to:

$$P_A+(1-P_A)(P_{A+1}+(1-P_{A+1})(P_{A+2}+(1-P_{A+2})(P_{A+3}+ (1-P_{A+3})(P_{A+4}+(1-P_{A+4})(P_{A+5}+ (1-P_{A+5})(P_{A+6}+(1-P_{A+6})(P_{A+7}+(1-P_{A+7}))))))) \qquad (4)$$

Values of the function P can be calculated for any trial values of a and b, and hence the proportion of licence records of women drivers which would be 'lost' after eight years. These estimates can be compared with the actual results from the analysis of the licence data; the optimal values of a and b are those which provide the best representation of these results.

Figure A1 shows the good agreement that was achieved for 'women licence-holders': similar results were achieved with 'women with provisional licences' and 'women with full licences'. This level of agreement suggests that (3) and (3') provide a suitable functional form for P, and hence a suitable basis for calculating the weights.

The weights are calculated directly from the function P. The loss function L(a,y) has the same form as (4), but the number of terms depends on the period over which the loss has occurred.

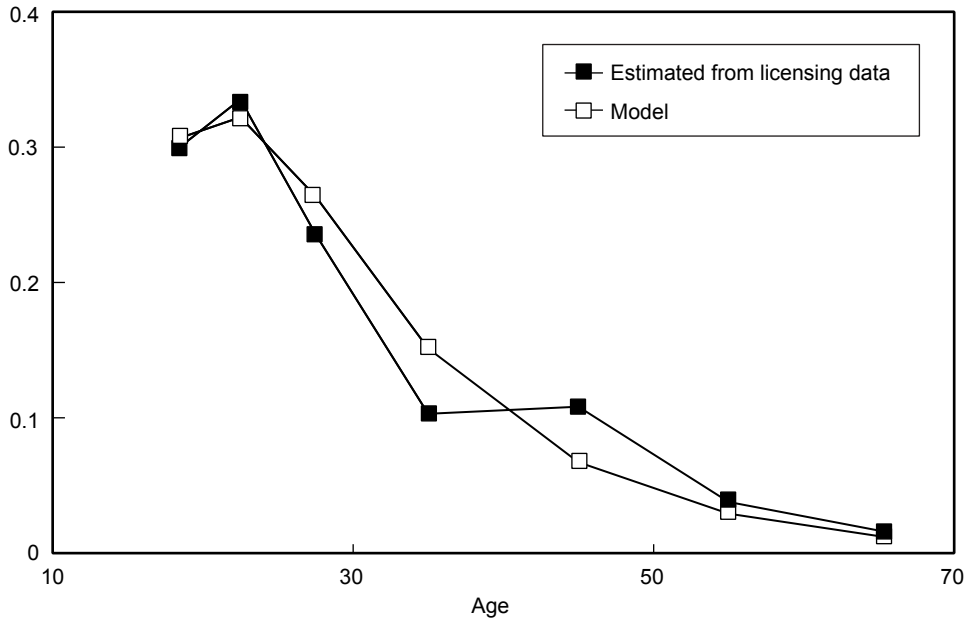


Figure A1 Proportion of women licence-holders 'lost' from sample within eight years

Abstract

The Driver Licence File maintained by the Driver and Vehicle Licensing Agency (DVLA) at Swansea has considerable potential for contributing to research into road safety. This report studies the incidence of driving offences by analysing the offence details contained in a sample of approximately 1 per cent of all licence records from 1985-1997.

The analyses show that the rate of driving offences committed varies considerably with the driver's age and sex, also by the type of licence held (provisional or full). In recent years, the rates have been highest among young men with full licences and lowest among women with provisional licences. There are clear regional differences, with the highest rates for men and women being found in North West England.

Corresponding analyses have also been made of three types of offence: Speeding, Careless Driving and Drink/Driving, although the limited sample size limits the precision of the results which can be achieved.

Related publications

TRL253 *Retesting as a penalty for dangerous driving* by L M Pearce. 1996 (price £25, code E)

TRL252 *Does retesting deter dangerous driving* by D G Harland and J Lester. 1997 (price £25, code E)

TRL232 *Drinking and driving in Great Britain - a review* by G Maycock. 1997 (price £25, code E)

PR111 *Cohort study of learner and novice drivers: Part 3 Accidents, offences and driving experience in the first three years of driving* by E Forsyth, G Maycock and B Sexton. 1995 (price £35, code J)

RR77 *Analysis of motoring offence details from DVLC driving licence records* by J Broughton. 1986 (price £15, code B)

CT42.2 Alcohol, drugs and driving update (1996-1998). *Current Topics in Transport: selected abstracts from TRL Library's database* (price £20)

CT80.1 Driver behaviour Update (1995-1997). *Current Topics in Transport: selected abstracts from TRL Library's database* (price £20)

CT84.1 Speed and road accidents Update (1994-1998). *Current Topics in Transport: selected abstracts from TRL Library's database* (price £20)

Prices current at July 1999

For further details of these and all other TRL publications, telephone Publication Sales on 01344 770783 or 770784, or visit TRL on the Internet at http://www.trl.co.uk.