Econometric Analysis of the Downward Trend in Road Fatalities since 1990

for Ministry of Transport

Prepared by Infometrics Ltd

July 2013

Abstract

An econometric analysis of the long term declining trend in the rate of road fatalities in New Zealand can explain more than 80% of the reduction. About 45% of the explained reduction is attributable to improvements in vehicle safety and to fewer motorcycle registrations. Another 19% is due to investment in more and better roads, while road safety advertising, breath testing and lower speed account for the remaining 36%. Without the historical changes in these variables that occurred since 1990 there would have been another 10,000 or so road deaths.

Table of Contents

This report has been prepared by Dr Adolf Stroombergen. Comments and corrections were provided by Dr Michael Keall of Otago University and Dr Charles Sullivan of Capital Research.

Email: adolfs@infometrics.co.nz

EXECUTIVE SUMMARY

Econometric analysis is applied to gain insight into the decline in the number of road deaths between 1990 and 2012. Over that period the number of deaths fell by 58% and the rate of fatalities per unit of travel by 74%.

While previous New Zealand studies have looked at the effect of particular safety initiatives such as advertising and random breath testing, to our knowledge none have attempted to identify the separate effects of the main fundamental contributing factors: the road, the vehicle and the driver.

Had there been no change in the number of fatalities per unit of travel since 1990, there would have been 12,300 more deaths between 1990 and 2012. The models estimated in this paper can explain approximately 10,000 of those as follows:

- 45% are attributed to the vehicle: improvements in vehicle crash worthiness and fewer motorcycles (which are in effect less safe vehicles) per unit of travel.
- 19% are attributed to the road, measured as real net investment in roading per unit of travel.
- 36% are attributed to the driver; a combination of the effects of advertising, breath testing, and average speed.

As with any econometric investigation a number of caveats should be considered. Firstly, error can arise through inaccurate measurement and faulty model specification, especially with regard to the handling of lagged relationships and correlation between explanatory variables.

Secondly, data scarcity has meant that some potentially important variables have been omitted from the models. For example the time between an accident and the arrival of medical care is known to be important to the chances of survival, and anecdotally that time is thought to have fallen. Unfortunately we have no consistent time series or suitable proxy to model such an effect. Changes in the mix of travel, such as urban versus rural could also affect the intensity of road deaths per unit of travel.

Thirdly even if the results are reliable they may not be applicable outside the sample period. For instance the proportion of motorcycle users accounted for by high-risk young males has declined, so the strength of the estimated relationship may be less in future. Conversely, there may be initiatives that will have more impact in the future – divided motorways or collision avoidance systems for example.

The success of past measures, be it investment in better roads, tighter frontal impact standards, breath testing and so on, should not be interpreted as automatically justifying them. All measures and initiatives should be subjected to thorough cost-benefit analysis before being continued, let alone expanded.

1. OVERVIEW OF NEW ZEALAND RESEARCH

We have not discovered any New Zealand studies that take an econometric approach to understanding the broad range of factors – the road, the vehicle and the driver – that could have contributed to the long term trend decline in road fatalities.

There are, however, numerous econometric studies that look particularly at the driver, through the effects of advertising and enforcement. Tay (2001) provides a summary of research before 2001 so there is little point in revisiting that earlier work. Instead we discuss Tay's findings. We also look at later studies by Guria and Leung (2004) and by Cameron and Sullivan (2011).

No research explicitly allows for the effect of changes in vehicle safety and the only research that allows for the effect of the road is a non-econometric study by Guria (1999),[1] who includes the safety component of road construction within a wider envelop of safety expenditure that also covers expenditure on enforcement, advertising and education. Guria finds very good benefit-cost ratios for the combined programmes.

In our overview below our interest is as much in the methodology of the studies as in the results.

Tay, Richard (2001): "Methodological issues in evaluation models: The New Zealand road safety advertising campaign revisited." *Road and Transport Research*, Vol 10, No 2, June 2001.

As noted, Tay focuses on advertising rather than the broader list of factors that affect the road toll. With regard to the Supplementary Road Safety Package (SRSP) – a policy initiative that was characterised by graphic television advertisements showing horrific accidents – Tay's summary shows that previous studies produced no consistent conclusions about the effectiveness of the campaign. He re-ran previous analyses, but paid more attention to autocorrelation than had been done before, and whether a trend variable should be included.

If a trend variable is included in an equation with trending time series, it will pick up the effect of missing variables, such as the effect of road quality improving over time. It may also, however, take some explanatory power away from other independent variables. A finding of statistical significance for other independent variables then provides some comfort that those variables do indeed have some explanatory power (assuming non-stationarity has been addressed). Of course the effects could be seriously understated.

Tay's results showed that more compulsory breath testing as part of the SRSP was not significant, but that the graphic advertising programme (as measured by Adstock) combined with breath testing did have a significant effect on the number of serious casualties during the first two years of the campaign.

Guria J. and J. Leung (2004): "An evaluation of a supplementary road safety package." *Accident Analysis and Prevention* 36 (2004) 893-904.

The study uses annual models and quarterly models to again estimate the effects of the Supplementary Road Safety Package (SRSP). In fact the study follows on from a

[1] Guria, J (1999): "An economic evaluation of incremental resources to road safety programmes in New Zealand." *Accident Analysis and Prevention* 31 (1999) 91-99.

previous one by the same authors together with Max Cameron in 2002[2]. That study raised a number of fairly serious concerns such as using GDP as a proxy for enforcement and advertising, ignoring autocorrelation and not considering short run versus long run effects.

In the newer study the equations are specified in semi-log format and so incorporate diminishing returns by assumption. The dependent variables are fatal crashes, fatalities and non-motorcycle fatalities, normalised by a travel volume index for the annual models. For the quarterly model the dependent variables have serious casualties added to fatalities.

Trend variables are included to pick up unspecified effects, but of course could also reduce the estimated coefficients on other explanatory variables. Dummy variables allow for the oil crisis, carless days and changes in the speed limit. And there are variables for new vehicle registrations, an estimated enforcement series, real advertising expenditure, compulsory breath tests, speed cameras, unemployment, and various interaction terms.

In the quarterly model some annual variables are assumed to be invariant over the year, implying some unusual quarterly changes.

Principal Components Analysis (PCA) is used to remove multicollinearity. This is usually a sensible strategy although it is worth noting that the lower variance in the estimates, which is the reason for using PCA could come at the cost of a (small) rise in bias. Not enough information is presented to tell whether even the first principal component has any sensible interpretation.

Unemployment emerges with a negative coefficient, but the theory as to why this should be is vague. Also its presence seems to confound the results for the safety policy variables.

The results produce significant estimated savings in fatalities due to the SRSP, but it was not possible to establish separate effects of the component parts – police enforcement and safety advertising.

Cameron M. & C. Sullivan (2011): *Review of Road Safety Advertising Programme in New Zealand*. Unpublished report to NZTA.

The report had a wider ambit than only evaluating the contribution of the advertising programme to road safety outcomes, but those other objectives are outside our scope.

The research used monthly data to analyse the effect of advertising recall on serious casualty crashes and hospitalisations of more than one day between 1998 to 2010. The estimated relationships were generally small and not statistically significant.

Measuring the effects of advertising is fraught with difficulty. Issues include the type of advertising, which media, how often, what audience, co-ordination with other advertising and enforcement activities, lagged effects, and so on. These are discussed by the authors.

A number of factors could have contributed to the absence of significant statistical links between advertising and serious accidents:

[2] Cameron, M. J. Guria & J. Leung (2002): *An evaluation of the supplementary road safety package July 1995 to June 2000.* Land Transport Safety Authority.

1. For a number of regressions alcohol crashes are confined to 'high alcohol hours' with speed related crashes confined to those that occurred outside those hours. This imposes an unnecessary restriction, ignores what is likely to be a substantial error margin on how accidents are classified, and ignores multiple accident causes.

2. Advertising recall was one the independent variables. As this is an intermediate variable situated between a more basic measure of advertising such as TARPs and the ultimate outcome variable – serious accidents or fatalities – it should have a stronger coefficient than a pure input variable. However, if it is measured with more error, as seems likely, the coefficient will be biased downwards.

3. One of the enforcement variables is the number of offences, albeit adjusted for the different frequency and amount of the fine associated with different types of offence. Unfortunately the number of offences has an ambiguous sign on the coefficient in the regression equation as changes in offences cannot be split between more enforcement and changes in behaviour.

4. Apart from advertising and enforcement variables the models include a travel volume index, GDP, unemployment, new vehicle registrations, new motorcycle registrations, a time trend, and dummy variables for monthly effects, for the change to the legal drinking age and for the change to the speed tolerance from a percentile to an absolute amount. Not surprisingly these variables generate severe multicollinearity, in spite of various combinations of variables being discarded (which is not always the best approach to dealing with the problem).

5. There is no allowance for lagged relationships or autocorrelation which, although no information on the behaviour of the residuals is presented, is likely to be a serious omission. (The importance of lagged relationships is discussed in section 2.2 below).

In defence of the authors they strongly recommend further econometric analysis of the relationship between advertising and serious accidents, which was beyond the scope of the study. Indeed as noted above this relationship was not even the prime focus of the study.

Reviewing overseas research has been beyond the ambit of this report, although some is cited in the following section in relation to particular points of interest. For a wider purview the reader is directed to the references in Keall et al.[3]

[3] Keall M., A. Stroombergen, C. Sullivan, M. Cameron and H. Weiss (2012): *An Analysis of Potential Factors behind the 2011 Reduction in New Zealand Road Fatalities.* Report to Ministry of Transport. From here on we refer to this report as KSSCW.

2. MODELLING LONG TERM EFFECTS

2.1 Approach

We begin our specification of a model by considering a simple identity equation for road fatalities:

$$F = F/T * T/V * V/P * P \quad (1)$$

F is fatalities (or serious injuries/crashes)

T is travel (VKT)

V is number of vehicles

P is population

Of course these are very aggregated variables. Travel for example is a heterogeneous amalgamation of many different journeys; commuting, delivery of goods, recreational travel and so on; occurring on roads with varying degrees of safety. Similarly a count of the population does not reflect differences in driving ability, nor does a count of vehicles capture differences in vehicle crash worthiness. However, none of this is necessarily problematic. If compositional variations lead to changes in fatalities (F) our task is to establish how. If fatalities are unaffected the variations are irrelevant.

Based on Equation (1) we could conceivably specify three models, one for each ratio, with explanatory variables as appropriate. In keeping with the accepted body of theory around road accidents, explanatory variables might be as listed in Table 1.

Table 1: Ratios and Main Variables of Influence

Dep Var	Independent Variables
F/T	road quality (surface, camber, median barriers, sealed shoulders etc), driver behaviour (speed choice, alcohol consumption etc), vehicle safety (affected by stability control, airbags, crumple zones etc) & vehicle type (motorcycle v car)
T/V	public transport cost & availability, vehicle $/km, state of economy, population age, vehicle occupancy, urbanisation or population density
V/P	public transport availability, $/vehicle, wage rates, state of economy, population age, urbanisation or population density

This schema is an approach to ameliorating high correlation between variables (multicollinearity) as the economic and demographic factors (for example) can be separated from factors that more directly affect the probability and severity of accidents. For example if the economic recession that has followed the global financial crisis has reduced how often people go out to drink, accidents involving drink-driving might fall.[4] This effect would be captured more in T/V, than in F/T. In contrast the effect of more enforcement would be primarily observable in F/T.

While the decomposition of influences could limit the explanatory power of each model, this should be a minor concern compared to the benefit of obtaining more reliable coefficients.

[4] See for example Lloyd, L., C. Reeves, J. Broughton and J. Scoons (2013): *Investigating the reduction in fatal accidents in Great Britain from 2007-2010*. Transport Research Laboratory.

For this research our interest is in F/T. Thus apart from a brief discussion about how the ratios T/V and V/P have changed over recent years, we do not seek to determine the direct effects of variables such as fuel prices and the state of the economy on the longer term accident trend. Their effects are either short term (as in KSSCW) or their long term effects are captured by other variables that directly affect F/T. For example strong GDP growth can be expected to deliver better quality roads while higher fuel prices persuade people to buy newer (more fuel efficient) cars, which also tend to be safer.

Figures 1-3 show the paths of the three ratios F/T, T/V and V/P. Note that:

1. Travel for 2012 is estimated on the basis of the first six months of 2012 and travel on state highways.
2. The number of vehicles for 2012 is estimated from light vehicle numbers for 2012.

Figures 1 and 2 depict steady downward trends, but the number of vehicles per person (Figure 3) seems to have peaked, although this could be temporary due to recent slow economic growth. One would not be confident in extrapolating either the recent trend or the longer term trend (given by the best fit line through the points).

Figure 1: Fatalities/Travel (number per billion km)

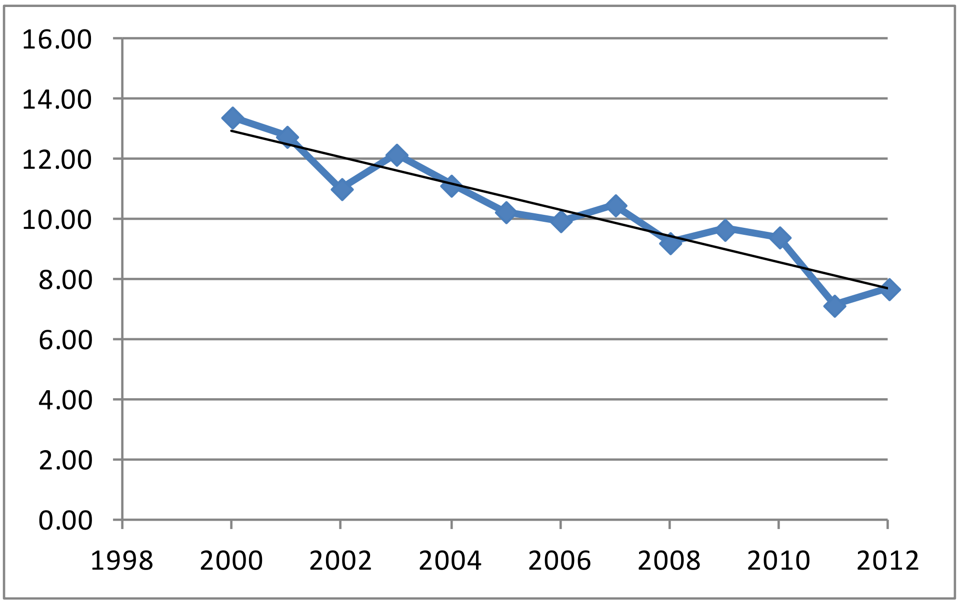


Table 2 decomposes the change in fatalities between 2000 and 2012.

Table 2: Decomposition of Change in Fatalities

Factor	Trend	Actual
	2000-2012	2010-2011
Population (P)	15%	1%
V/P	9%	-1%
T/V	-10%	0%
F/T	-41%	-24%
Fatalities (F)	-32%[5]	-24%

[5] Check: 1.15*1.09*0.90*0.59≈0.68

Figure 2: Travel / Vehicles (km per vehicle)

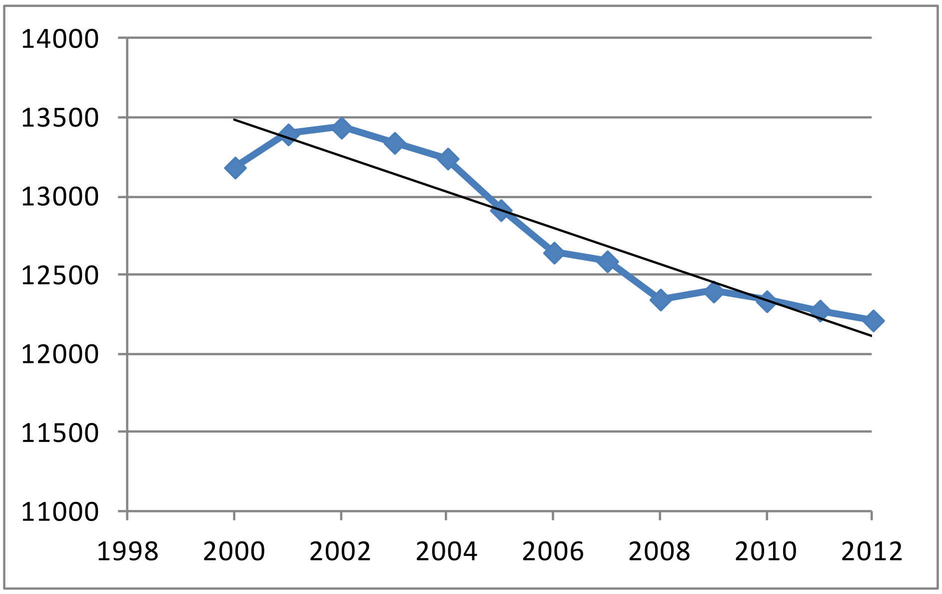
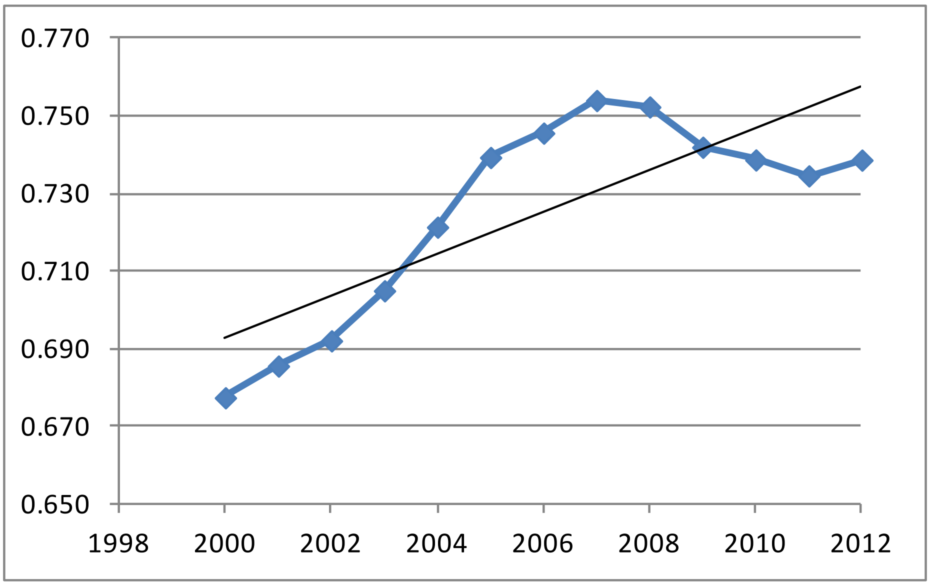


Figure 3: Vehicles / Population (number per person)



The contribution of lower F/T to the declining road toll since 2000 is clear. Less travel per vehicle also contributed, but this was more than offset by higher vehicle ownership and a growing population.

In context of KSSCW which looked at the large drop in fatalities between 2010 and 2011, the changes in population and V/P offset each other, and with no change in T/V the drop in fatalities matches the drop in F/T.

It is necessary when reducing the road injury rate as expressed in equation (1) to lower at least one of the component ratios F/T, T/V, and V/P. Most people would probably share the view that limiting the growth of the population in order to reduce the number of road fatalities would be silly. Similarly with regard to the number of vehicles per person, although there may be other reasons to discourage the expansion of the vehicle fleet. Going a stage further, road safety policies are not generally targeted at reducing the amount of travel per vehicle. That is usually the target of policies to curtail greenhouse gas emissions or to lessen congestion.

Thus our focus for econometric estimation will be entirely on F/T – because it is the factor that has the largest effect on fatalities; because it is the factor that is the target of road safety interventions, whether driver education, enforcement, or better road engineering; and because it is the main measure showing how improvements in vehicle safety affect the road toll.

Also, F/T is the only series for which the underlying data can be readily extended back – to 1989 in fact. Data on the number of vehicles begins in 2000, thereby precluding longer term analysis of T/V and V/P.

The variables listed in Table 1 as potential explanatory variables for F/T have long term effects. Improvements in vehicle safety take many years to percolate through the vehicle fleet. Better roads cannot instantly be seen to lead to fewer fatalities – it takes time for the effects to be noticeable above random variation.

Some variables may also have short term effects, although we know from KSSCW that quarterly variation in Advertising Recall, TARPs, CBT and MBT had no power to explain the departure of fatalities from a medium/long term (since 2000) trend.

To capture long term effects we need lagged explanatory variables. Unfortunately this frequently generates severe multicollinearity and also uses up valuable degrees of freedom. To circumvent this situation the coefficients on lags are frequently approximated by some polynomial function which allows the lagged effects to be captured as a lag of the dependent variable. This is the procedure we adopt here. Appendix A provides more detail.

2.2 Model Variables

In this section we look at how F/T has changed since 1989 and construct a number of explanatory variables to use in the econometric analysis. As improvements in roads and vehicles reduce the risk of accidents occurring and, given an accident, reduce the risk of serious injury and death; our aim is to have at least one variable that captures the long term effect of each of the road, the driver and the vehicle.

F/T since 1989

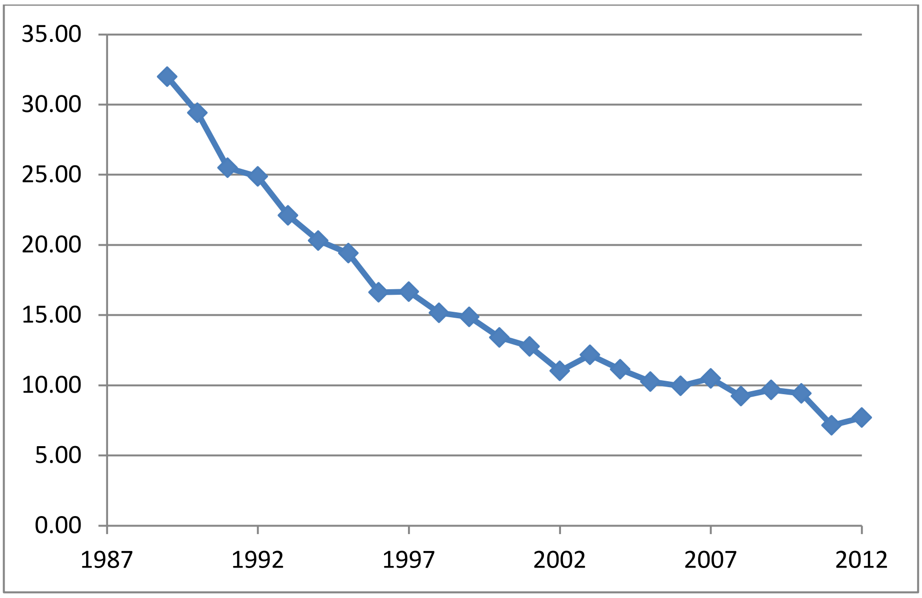
Figure 4 shows F/T going back to 1989, with travel before 2000 estimated by splicing the state highway travel series onto the newer series from the Vehicle Fleet database. For the overlap period the mean absolute error is 0.7%. In any case we are more interested in the general trend than in precise short term alignment.

The decline in the extended series is less linear than in Figure 1, showing a more convex shape, which makes sense as a linear decline would rapidly approach infeasible numbers. Since 1989 the fatality rate has declined by over 75%. Had the rate remained as it was in 1990,[6] the number of fatalities in 2012 would have been about 1177 compared to the actual number of 308. Total fatalities over the period 1990-2012

[6] We do not begin with 1989 in order to facilitate later comparison with the econometric models which 'lose' the first observation through the inclusion of a lagged term in the estimation.

would have been approximately 23,360 compared to an actual 11,004 – a difference of over 12,300.

Figure 4: Fatalities/Travel (number per billion km)



The Vehicle (Crash Worthiness)

Vehicle safety has two sets of features: primary features, such as ABS and traction control that help avoid an accident, and secondary features such as (front and side) airbags and frontal impact standards that protect occupants in the event of an accident.

Newstead et al[7] estimate the probability of a serious crash given an injury crash (and also the probability of an injury crash given a crash), by vehicle year of manufacture (from 1964 to 2009). We can use their results as a proxy measure of vehicle safety – some makes and models came with stability control before others and vehicle safety regulations differ by country of manufacture.

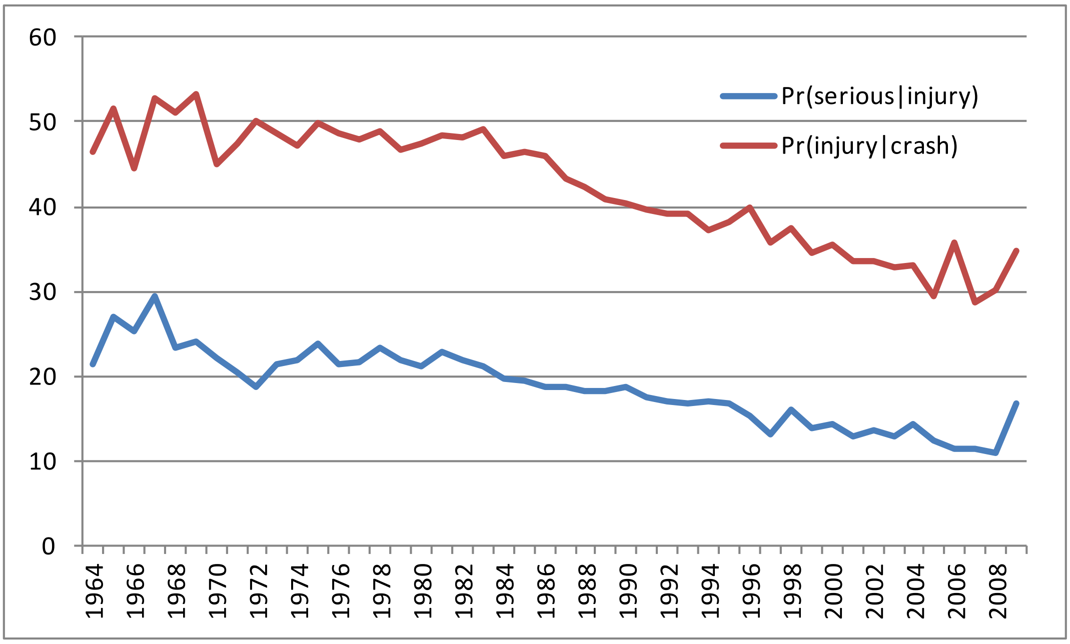
The probabilities are estimated by fitting logistic models of driver injury against a number of explanatory variables: year of manufacture, year of crash, speed zone, the number of vehicles in the crash, age and sex of the driver. Although few regression statistics are reported the detailed discussion of the data issues and modelling process suggests a robust set of results.

Although the models have few explanatory variables this is not necessarily a concern. Even in a regression with the year of manufacture as the only explanatory variable, its coefficient would be unbiased if there is no interaction between it and other explanatory variables. The models do allow for yearly fixed effects, thereby obviating the influence of other factors such as enforcement programmes. As may be seen in Figure 5, most of the improvement in crash worthiness (decline in accident probabilities) begins in the early 1980s.

---

[7] Newstead S., L. Watson & M. Cameron (2011): *Trends in Crashworthiness of the New Zealand Vehicle Fleet by Year of Manufacture: 1964 to 2009.* Supplement to Report 304, Vehicle safety ratings estimated from police reported crash: 2011 Update. Monash University.

Figure 5: Crash Worthiness
(%)



Weighting the probabilities of a serious injury accident, given an injury accident, by the age composition of the vehicle fleet in each year, gives the probability of a serious injury accident, given an injury accident, by year. This is illustrated in Figure 6. However, a number of caveats should be noted:

1. The data for 2000 to 2008 is robust as it is based directly on Newstead et al's crash worthiness data and the age composition of the vehicle fleet, obtained from the Ministry of Transport's (MoT) vehicle fleet model. Although the former series extends to 2009, Newstead et al suspect its validity owing to a small sample size.

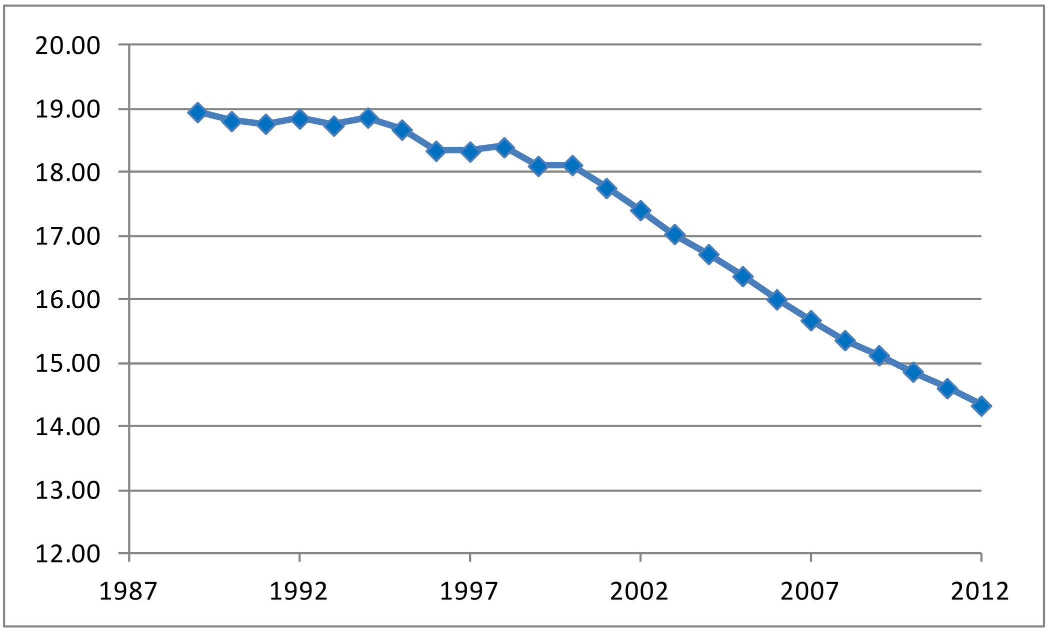
2. From 2009 to 2012 the estimated series still uses the MoT data on the age composition of the vehicle fleet, but the crash worthiness data has been extrapolated from 2008 to 2012 using a linear trend, which as suggested by Figure 5 seems to be reasonable, but could be wrong.

3. From 1989 to 1999 the estimated series uses the crash worthiness data from Newstead et al, but the age composition of the vehicle fleet is proxied by the average age of the fleet.[8]

As can be seen in Figure 6 the probability of a serious injury, given an injury, averaged over the whole fleet, does not begin a significant annual decline until about 2000, reflecting the age of the vehicle fleet. From 2000 the rate of decline is about one third of a percentage point each year.

[8] Obtained from the Motor Industry Association.

Figure 6: Probability of Serious Crash Given Injury Crash (%)



Due to the small number of serious accidents there is a large amount of randomness around the ratio of fatalities to serious accidents. Small differences in the angle of collision impact, the speed of the impact, and the swiftness and appropriateness of the driver's reactions can literally mean the difference between life and death. Thus a marked decline in the probability of a serious injury crash, given an injury crash, should go a long way to explaining the reduction in F/T over the sample period. As shown in Stroombergen & Keall (2013),[9] over time serious casualties and fatalities track together.

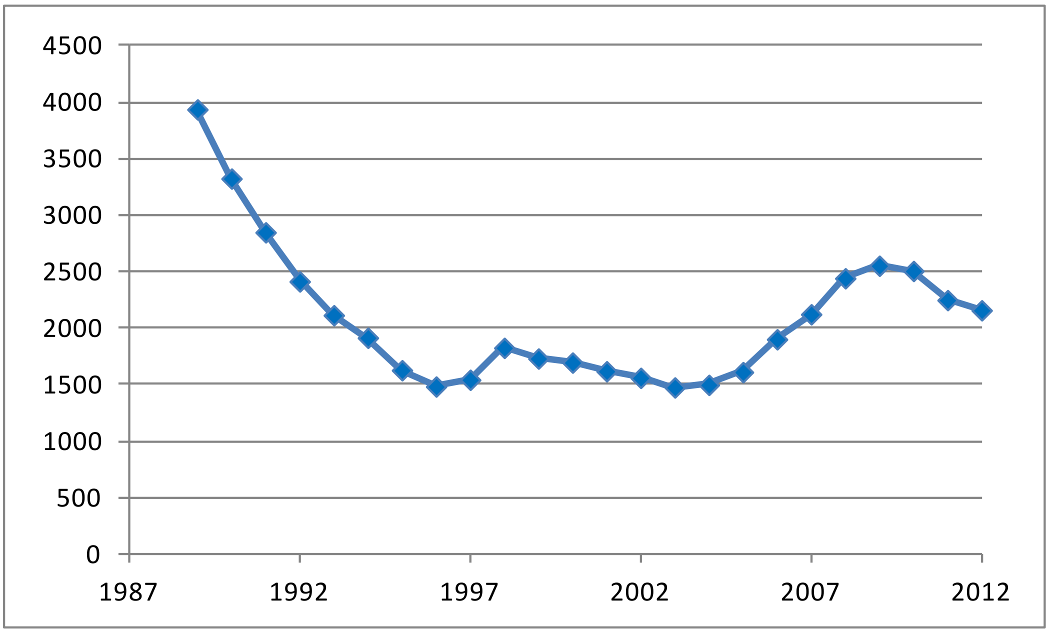
The Vehicle (Motorcycles)

From the results in KSSCW we know that changes in motorcycle registrations help to explain changes in fatalities on a quarter to quarter basis. A longer term effect is also plausible if one thinks of motorcycles as analogous to a vehicle safety issue. A motorcycle is essentially a vehicle with very little protection for the driver. Thus we would expect the overall safety of the vehicle fleet from a passenger protection perspective to be inversely correlated with the proportion of motorcycles in the fleet. Accordingly we use motorcycle registrations as another measures of vehicle safety in our model.

To maintain consistency with the dependent variable we divide motorcycle registrations by the volume of travel. Figure 7 shows the resultant series.

[9] Stroombergen, A. & M. Keall (2013): *Further Econometric Investigation of the 2011 Reduction in Road Fatalities.* Report to Ministry of Transport.

Figure 7: Motorcycle Registrations per Unit of Travel (No. per billion km)



The Road (Expenditure on Infrastructure)

The safety quality of road infrastructure is difficult to measure. The New Zealand Road Assessment Programme (KiwiRAP) measures roads by a safety rating, but as it is partly based on fatal crashes it cannot be described as an input measure into an econometric model that is independent of the number or rate of fatalities.

There is some data on road roughness and skid resistance, but not enough for econometric analysis. The various measures of pavement condition produced by NZTA are available only from 2000. In any case road condition is but one characteristic of road safety. Other characteristics include camber, lane width, lane separation, lighting and roadside hazards such as power poles, culverts and loose gravel. In the absence of such measures we will use real expenditure on road construction, renewal and maintenance as a measure of road quality. However, there are several issues in adopting this measure which should be noted:

1. Diminishing returns can be expected if projects have been ranked and undertaken in order of benefit-cost ratios, such as beginning with accident black-spots. Still, as safety is not the only component on the benefit side of the equation (with travel time being the other main component), it seems unlikely that diminishing safety returns would be significant.

2. Even if safety was by far the dominant type of benefit, diminishing returns would depend on an accurate evaluation of the safety effects of all of the road characteristics mentioned above, followed by a strict order of project implementation. Again this seems unlikely.

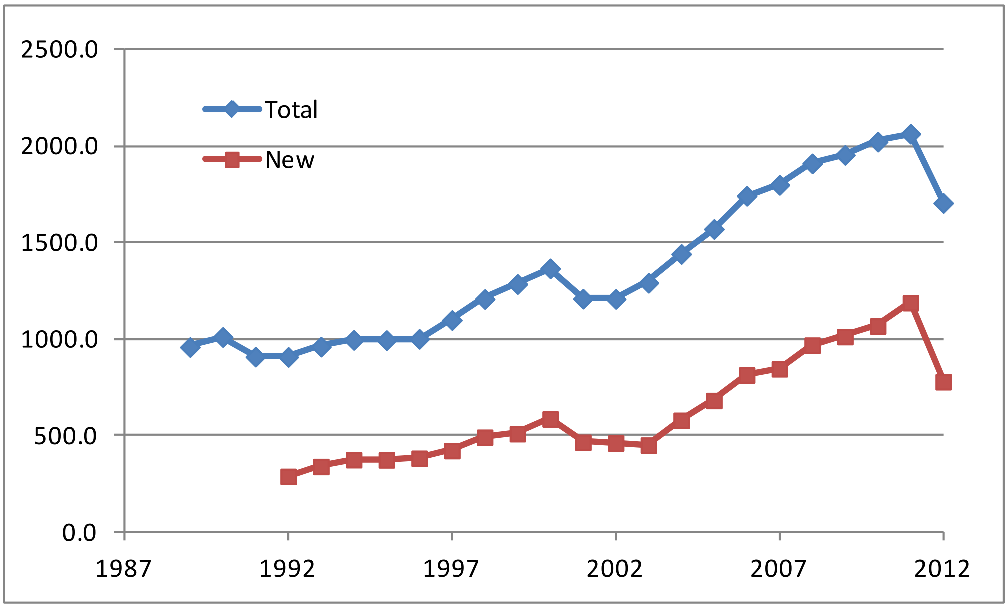
3. Relative benefit–cost ratios are likely to vary over time in response to changing travel patterns. Guria (1999)[10] makes the point that ex ante estimates of the benefits of roading expenditure can differ substantially from ex post benefits.

4. Casual observation of New Zealand roads suggests that the country is nowhere near the point where extra expenditure on road safety will deliver only marginal benefits (assuming at least reasonably sensible targeting of such expenditure).

The preceding points mean that while we still expect expenditure on roads to have a generally positive impact on road safety, especially over periods of years, short term impacts are subject to considerable noise.

Figure 8 shows total real expenditure on roads (new construction, renewal and maintenance) and real expenditure on new construction only, by NZTA and local authorities, deflated by the PPI-Output series for Heavy and Civil Engineering Construction.[11]

Figure 8: Real Expenditure on Roading
($m, Dec 2010 prices, years ended June)



As it is the stock, not the flow of investment in roading that affects road quality we construct a simple capital stock series, making the assumption that expenditure on maintenance and renewals is exactly that – that is, it covers depreciation. Clearly this will not be true every year, but as our interest is in the trend effect this limitation is acceptable. In any case we would not expect short term variation in road quality (even with an ideal measure) to correlate strongly with short term variation in fatalities.

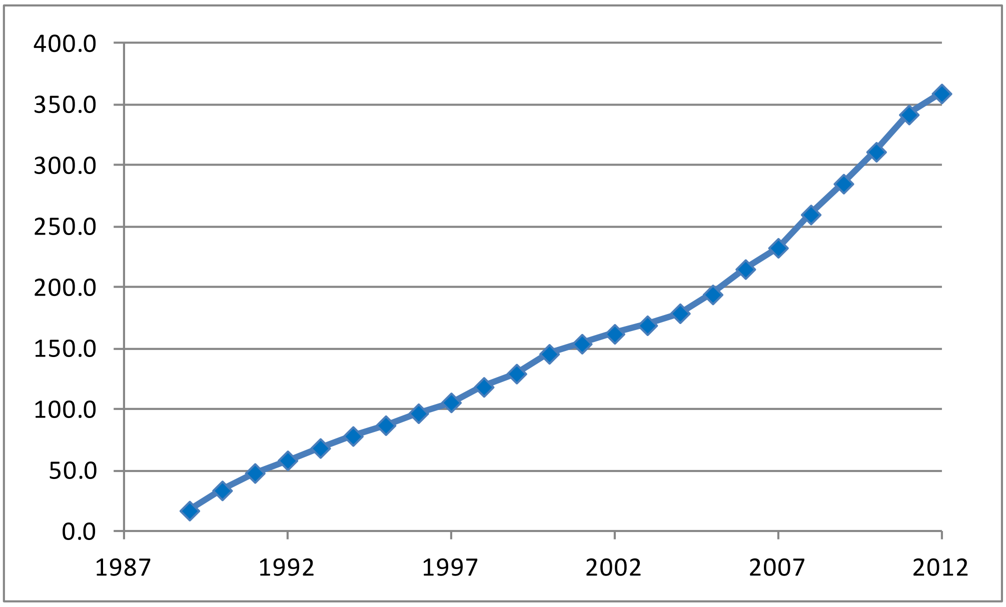
Note that any investment prior to 1989 is ignored so the series does not correspond to the actual value of New Zealand's roading capital stock. For our model, however, this is immaterial.

The capital stock series is divided by travel in order to make it consistent with the dependent variable (F/T). Figure 9 shows the result. Data for 1989-1991 is estimated

[10] *op cit*
[11] To deflate expenditure before June 1994 we use the Capital Goods Price Index for Transport Ways.

by splicing an older series from the former Land Transport Safety Authority onto the Ministry of Transport's new series.[12]

Figure 9: Real Roading Capital Stock from 1989 per Unit of Travel ($m per billion km)



The Driver (Advertising and Enforcement)

One might hypothesise that the safety of the driver is related to common demographic and economic variables such as age, sex, education and income. The sign of relationships, however, can be difficult to deduce theoretically. The effect of age for example might be quadratic – driver safety initially improving with experience, but then deteriorating as age reduces reaction times and increases fragility in crashes. Does higher income lead to safer driving or does it just enable the purchase of newer, and thus safer vehicles?

What makes people safer drivers? Education, advertising and enforcement should theoretically have a beneficial effect (in most cases), but not all measures that increase safety are costless to the individual. For example driving more slowly increases travel time, but travel time is generally valuable.

Advertising and Enforcement

The relationships of both advertising and enforcement to desired changes in driver behaviour can be difficult to measure reliably. First, they may interact; for example, a substantial increase in enforcement may have wider effects if advertising increases the number of drivers who are aware that the probability of getting caught has increased. Second, it can be difficult to predict the relevant timeframe for effects – how quickly drivers may respond, and how durable any effects might be. Third (and relatedly), both may contribute to long-term effects (such as the establishment of changed social norms) that are different from short-term effects, such as an increase in the perceived probability of being caught for drink-driving.

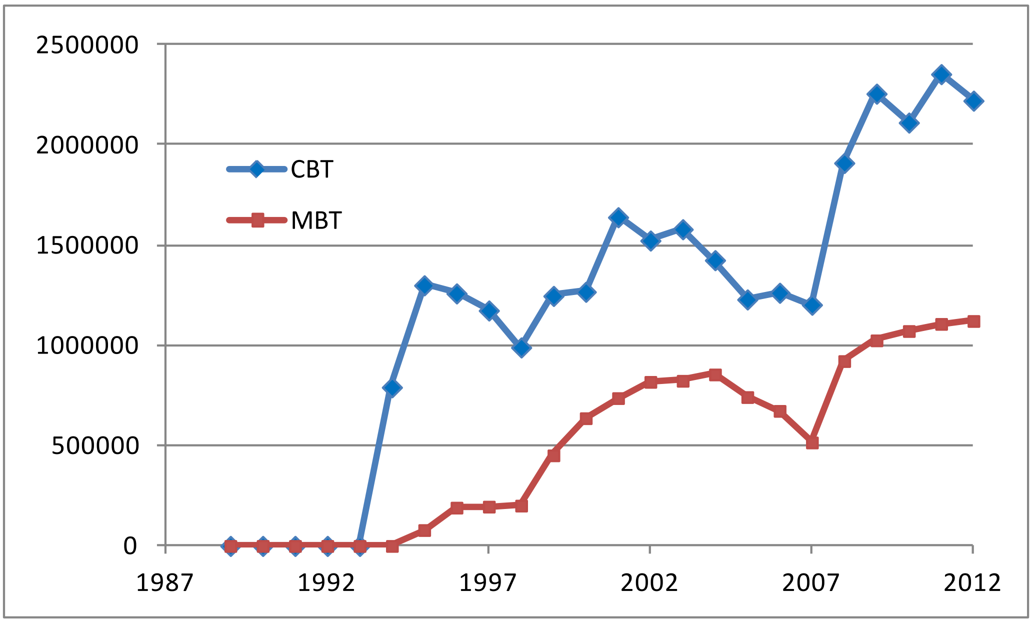
[12] http://www.transport.govt.nz/ourwork/TMIF/Pages/II015.aspx

An issue with enforcement is that while it may be effective in addressing the primary target (for example reducing speeds), that success does not necessarily translate into an effect on a less direct target – fewer fatalities. This is especially likely in the short term where the effect is via deterrence (with regard to the primary target variable) as opposed to longer term effects that may establish new behavioural norms.

For enforcement we use the number of compulsory breath tests (CBTs) and the number mobile breath tests (MBTs). They are pure input measures. Data for these series begins around 1994, but prior to then such testing did not occur or was very rare. We do not use drink-drive infringements, or indeed any infringements data as an explanatory variable, as it is impossible to tell whether a change in convictions is due to a change in inputs (enforcement) or to a change in driver behaviour; with opposite signs on the rate of fatalities.

Figure 10 shows CBT and MBT. Note that the figures for the last quarter of 2006 and the first three quarters of 2007 are known to be understated, but the correct values are unknown.[13] In the models estimated in the next section we test for the significance of the incorrect data with a binary variable.

Figure 10: Number of CBT and MBT



For advertising the only variable with a long enough period of coverage is Target Audience Rating Points (TARPs). It is not a pure input variable as it (ostensibly) measures the extent to which advertising hits the desired audience. However, it may be better than a pure input variable, as it is not obvious what such a variable would actually look like. Measures such as real spending, broadcast minutes (for television and radio) or column space (for print media) are fairly crude as they take no account of the type and content of advertising or when it is broadcast/seen. While TARPs is not a perfect metric it does at least provide a measure of quality that should be one step up from counts of dollars or minutes.

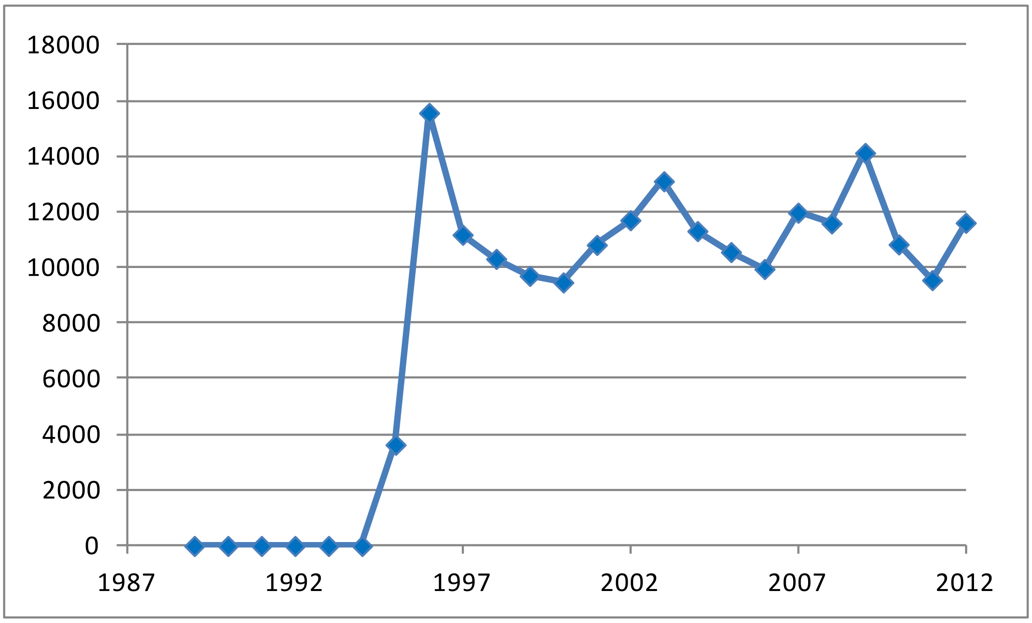
Another commonly used series to measure advertising is Adstock, which is a function of TARPs and an estimated or assumed 'half life' of advertising.' However, as discussed in

[13] See KSSCW.

Appendix A, our estimated equations implicitly incorporate a half life type of effect so there is nothing to be gained by adopting an Adstock measure.

Figure 11 shows the TARPs series. There is no data prior to 1995 although we understand that television advertising directed at road safety was rare before then.

Figure 11: Number of Target Audience Rating Points



The Driver (Speed)

A reduction in average speed is likely to be correlated with fewer fatalities. So ideally we would like to know whether changes in speed enforcement have contributed to the decline in fatalities. However, to our knowledge these is no good series that measures speed enforcement. Guria and Leung (2004)[14] used estimated enforcement hours, but the series was conceptually weak and cannot be updated. And as noted above we do not use speeding infringements as an explanatory variable as it is an output, not an input, causing its sign to be ambiguous.

We do have data on observed mean rural speed, though clearly changes in the series cannot be attributed solely to changes in enforcement; safety promotions may also contribute.

Figure 12 shows mean rural speeds compiled from Ministry of Transport data. From 1988 to 2004 data on speeds was collected from open road winter surveys. Since 1995 speed data has been obtained from police district surveys, but there is a sufficiently good overlap between the two series – in terms of both length and profile – to enable the new series to be extended back to 1988.

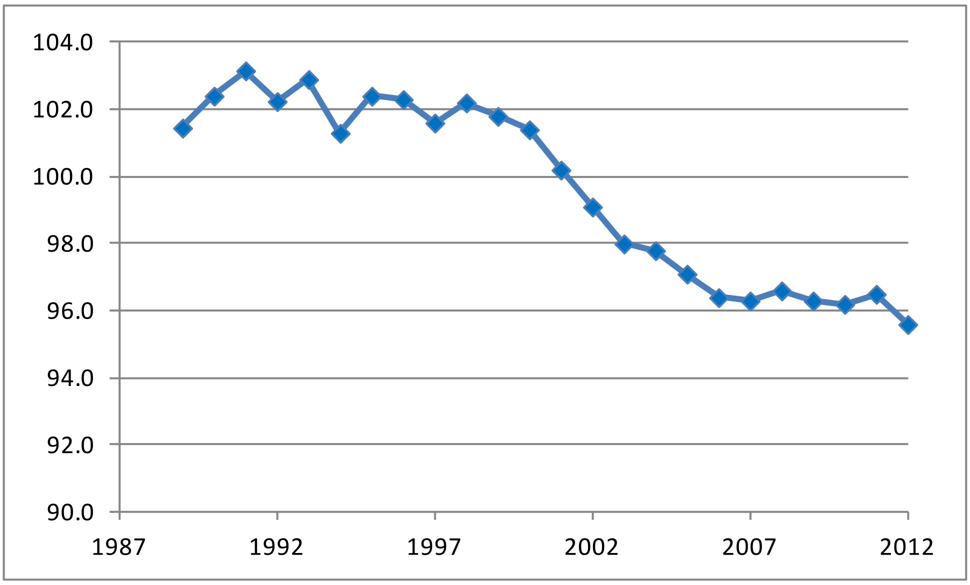
Mean speeds did not change much between 1988 and 1999. The next seven years saw a steady decline, but there has been little change since 2006. Although the overall slope of the trend is consistent with the decline in F/T, the reductions in F/T since 2006 and

[14] *op cit*

before 1999 occurred without any change in mean speeds. Thus we should not expect to see the change in rural speed as a major explanation of the long term decline in F/T.

Also, the shape of the curve up to 2006 is similar to that for vehicle crash worthiness shown in Figure 6, suggesting a potential problem with multicollinearity.

Figure 12: Mean Rural Speed (km/h)



3. ECONOMETRIC ANALYSIS

In accordance with the discussion above we explore two slightly different model specifications; a model incorporating a lagged dependent variable and an Error Correction Model (ECM).

The equations are specified in levels rather than as a log-linear (Poisson) regression that is sometimes appropriate when the dependent variable has only integer values, such as the number of fatalities. The F/T series is not confined to integer values.

Both models include variables for the road, the vehicle and the driver, as detailed above. Also tested were dummy variables for the lowering of the legal drinking age in December 1999, and for the change in speeding tolerance from a percentile figure to an absolute km/hr figure in July 2000. The latter is completely insignificant which is not surprising given the small effect that the change is likely to have had on driver behaviour.

The coefficient on the drinking age dummy is significant but, perhaps contrary to expectations, is negatively signed. However Stillman and Boes (2013) find:

> "no evidence for an increase in alcohol related road accidents at the time of the law change for teenagers relative to slightly older drivers." [15]

This suggests that the coefficient should be close to zero. The presence of this variable also reduces the significance of the breath test variable. Thus rather than have a variable with a sign that is not easily explained, which may not be robust, and which interferes with a variable that has a plausible sign (as discussed below) it is omitted from the final equation specification.

3.1 Model 1: Lagged Dependent Variable

A lagged dependent variable model was estimated in three ways; firstly with Ordinary Least Squares, then with Cochrane-Orcutt estimation to accommodate serial correlation and finally using ridge regression to mitigate the problem of multicollinearity. Table 3 shows the results of the first and third procedures.

Table 3: Lagged Dependent Variable Model

	OLS estimation		Ridge regression		
	Coeff	p-value	Coeff	p-value	Effect over 23 years
Lagged F/T	0.670	0.007	0.261	0.000	
Road investment/travel	0.0007	0.979	-0.0123	0.000	1855
Vehicle crash worthiness	1.314	0.495	0.538	0.000	1015
Motorcycle registrations/travel	0.0014	0.204	0.0023	0.000	3287
Total breath tests (CBT+MBT)	-0.250	0.539	-0.532	0.050	1073
TARPs	-0.209	0.846	-0.164	0.010	1596
Mean rural speed	-0.336	0.531	0.277	0.010	928
					9753
R^2	0.983		0.974		
Sum of squared errors	14.54		21.53		
Autocorrelation test (LM test)[16]	fail		pass		
Normality test (JB test)	pass		pass		

[15] http://cms.mngt.waikato.ac.nz/Portals/17/StevenStillman_21.03.13.pdf

[16] The Durbin-Watson statistic for autocorrelation is biased in the presence of a lagged dependent variable. Thus we use the Lagrange Multiplier test. This also tests for higher order autocorrelation.

The first equation shows a typical picture for a model plagued by autocorrelation and multicollinearity.[17] Apart from the lagged dependent variable no other variable is even remotely significant.[18] A number of series share common trends and some have little variation apart from their trend.

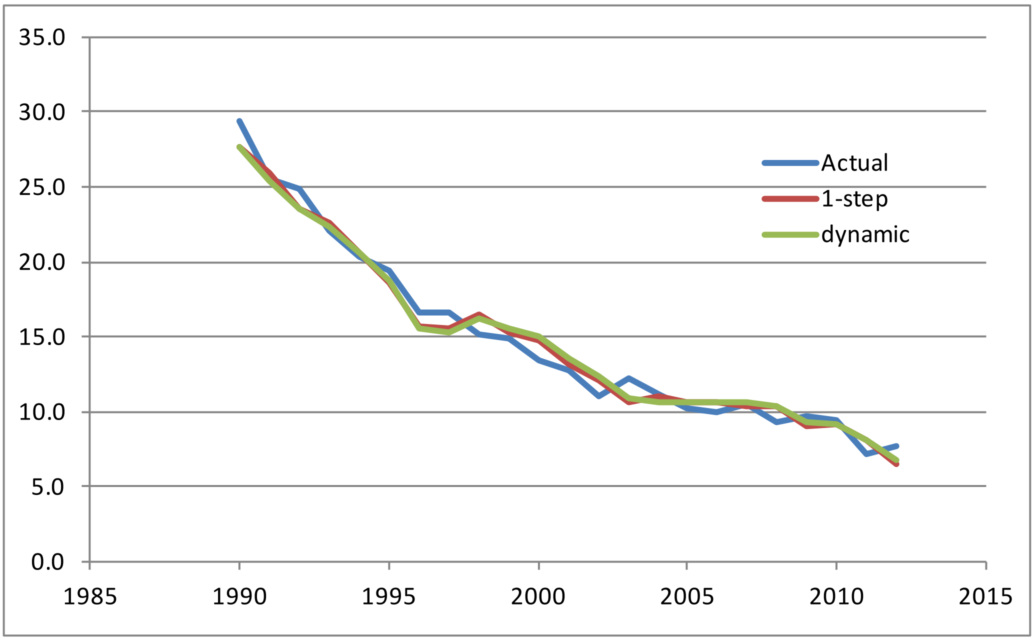
A dummy variable for the undercount in the breath test (CBT+MBT) series during 2006/07 was tested and subsequently omitted.

The ridge regression estimator reduces multicollinearity by shrinking the estimated coefficients towards zero. The shrinkage parameter (k) is 15%. This introduces bias into the coefficients, but the variances are much lower, producing estimates that are more reliable overall. The choice of the value of k is based on the following criteria:

- Stability of the coefficients as k changes.
- Coefficients not having implausible values.
- The sum of squared errors not rising to an unreasonable value, or be large in relation to the processes generating the data.

Figure 12 shows the actual and estimated values for F/T.

Figure 12: Model 1 Results for F/T (number per billion km)



The estimated equation is consistent with a number of different structural specifications, including partial adjustment and/or adaptive expectations models – refer Appendix A. As the ridge estimator has removed the serial correlation (including higher order serial correlation), we can infer that the model is dominated by partial adjustment processes

[17] The standard condition number tests on the eigenvalues of the correlation matrix indicate severe multicollinearity.

[18] We assume a two-tailed test, although one could argue that a one-tailed test would be more appropriate if the sign of a coefficient is theoretically unambiguous.

and lagged effects. However, rather than solve for the implied effect of each explanatory variable, it is more informative to calculate the contribution that each variable has had to the declining number of fatalities over the sample period. These effects are presented in the last column of Table 3. They show the impact on the number of fatalities if each explanatory variable had not changed from its initial value in 1990, given the actual amount of travel in each year.

To put the numbers in perspective; the total number of fatalities from 1990-2012 was 11,004. The model's one-step-ahead estimate for the whole period (that is using the actual value of the lagged dependent variable each time a forecast value is calculated) is 11,001 fatalities. The model's dynamic estimate (where only the 1989 actual value is used) is slightly higher at 11,014 fatalities, while the sum of absolute errors from the dynamic process is 640, an average error of 28 fatalities per annum.

Using the dynamic formulation the last column in Table 3 shows that if none of the explanatory variables had changed in value since 1990, the total number of fatalities over the period 1990-2012 would have been higher by 9753 – about 90% more than what actually occurred, or an extra 424 deaths per annum on average. That this number is not as high as the 12,300 referred to in Section 2.2 is due to model error and because the lagged term in the equation means that the model converges to a value for F/T that is different from the starting value. In less mathematical parlance: holding the explanatory variables constant at their 1990 values does not mean that F/T remains at its 1990 value because lagged effects, such as from spending on road infrastructure continue to generate benefits beyond the year in which they occur.

The largest contribution at 3287 fatalities comes from the decline in motorcycle registrations. The actual number of motorcycle fatalities from 1990-2012 was 1214. In the past young males were the main motorcycle users, so an already high-risk group used the highest risk vehicle. However, the proportion of motorcycle users accounted for by this age group has declined, so the strength of the estimated relationship between registrations and F/V may be less in future.

Together with the effect of the improvements in vehicle crash worthiness, about 44% of the decline in fatalities since 1990 can be attributed to 'the vehicle'. The road accounts for 19% while driver-related factors (encompassing advertising, breath testing and speed) account for the remaining 37%.

It is interesting that the mean speed and advertising variables are both significant, suggesting that the overlap between the two variables is not that strong. That is, the reduction in mean speed cannot be solely attributed to advertising campaigns, which of course do not target just speed. The implication is that other factors such as enforcement and education have had an effect.

The marginal significance on breath testing may indicate that while such enforcement is perhaps effective in lowering drink driving in cities (which is where most breath testing occurs) fatalities are relatively rare in city road accidents.

Of course all estimated coefficients contain error. Table 4 expresses each variable's standard error as a proportion of the coefficient value. The lag coefficient is the strongest, with a proportional error of about 10%.

The weakest coefficient, not surprisingly given its p-value in Table 3, is that on total breath tests (TBT), where the proportionate error is approaching 50%. A 95% confidence interval based on a t-distribution (which is appropriate here) would multiply each standard error by a factor of 2.1, which means that the upper bound on the (negative) breath tests coefficient is only just below zero.

All of the other coefficients have 95% confidence intervals with lower or upper bounds that are well away from zero, with road investment, vehicle safety, and the lagged effect being strongly non-zero.

Table 4: Standard Errors

Variable	se as % of coeff
Lagged F/T	11.8
Road investment/travel	17.9
Vehicle crash worthiness	18.7
Motorcycle registrations/travel	20.5
Total breath tests (CBT+MBT)	47.0
TARPs	34.1
Mean speed	34.1

It is worth reiterating that the ridge regression estimator biases the coefficients towards zero, meaning that the estimated coefficients are unlikely to be substantially over-estimated, unless the series are highly correlated with an omitted series.

Also, a good fit to past data does not mean that the model will necessarily apply beyond the sample period. For example there may be an eventual saturation effect from advertising. Furthermore the historical effects shown in the last column of Table 3 are the result of each coefficient multiplied by the change in the explanatory variable to which it applies. If the rate of increase in vehicle safety or investment in better roads begins to decline, so will the rate of decline in F/T, even if the coefficients themselves remain applicable.

3.2 Model 2: Error Correction Model

The fundamental theory behind a co-integrating relationship is that a linear combination of two (or more) non-stationary series is stationary, implying that any divergence of one series from the other is temporary. In this case an Error Correction Model (ECM) consisting of an equation that describes the long run relationship and another equation that describes the temporary departures from the long term relationship, is an appropriate model. However, an ECM can still be a valid and useful model when co-integration does not prevail.

Twenty-three observations is not really large enough for robust statistical testing for co-integration, although 23 years is a long enough period in this context for a long run relationship to apply between the rate of fatalities and the factors that affect it.

Table 5 shows the results of re-specifying the above model as an ECM. In the long term equation all of the variables are significant but, as expected without the inclusion of lags, the model suffers from severe autocorrelation. The residuals, however, are Normally distributed and are stationary.

Although the focus of this study is on long term relationships, we need to look at the error correction equation in order to allow for lagged effects. It makes sense then to also consider other variables that might have additional effects.

We have discussed issues around variables such as advertising and enforcement. These variables are likely to have both long term and short term effects. In KSSCW no short term quarterly effects on fatalities were ascertained for changes in these variables, but this does not preclude effects from annual variations – which of course further contribute to the long term effects over many years as well. Consequently it makes sense to incorporate the first difference of these variables into the error correction component of the ECM.

Table 5: Error Correction Model

	Long term eqn ridge regression (k=0.15)		Error correction equation		
	Coeff	p-value	Coeff	p-value	Effect over 23 years
Road investment/travel	-0.0183	0.000			1822
Vehicle crash worthiness	0.798	0.000			978
Motorcycle registrations/travel	0.0037	0.000			3765
Total breath tests (CBT+MBT)	-0.691	0.025			949
TARPs	-0.234	0.002			1650
Mean rural speed	0.395	0.003			876
					10041
Lagged residual from LT eqn			-0.730	0.001	
Change in CBT			-0.902	0.041	93
Change in TARPs			-0.117	0.110	
R^2	0.970		0.471		
Sum of squared errors	34.2		18.5		
Autocorrelation test	fail (DW)		pass (LM)		
Normality test (JB test)	pass		pass		

Testing revealed that the change in TARPs is very significant with the expected negative sign. The change in CBT has marginally higher significance than the change in total breath tests (TBT=CBT+MBT), so it is included in the equation. The change in MBT is not significant.

As in KSSCW we also looked at the change in real petrol prices as a possible short term explanatory variable. Not surprisingly it emerges insignificant here as its predominant effect on fatalities is likely to be via the demand for travel, not via F/T. Also, factors that help explain short term volatility in the road toll between 2000 and 2010 (in KSSCW) may not be the same as those that acted over the longer period between 1990 and 2000. Petrol prices in particular were much more stable during that decade and so would not have had the effect on travel that occurred when prices changed rapidly.

The most important variable in the short term model is the lagged error term from the long term model. It has the expected strong negative sign (-0.73), meaning that any divergence between F/T and the contemporaneous effect of the explanatory variables is quickly closed through the lagged effects of those variables.

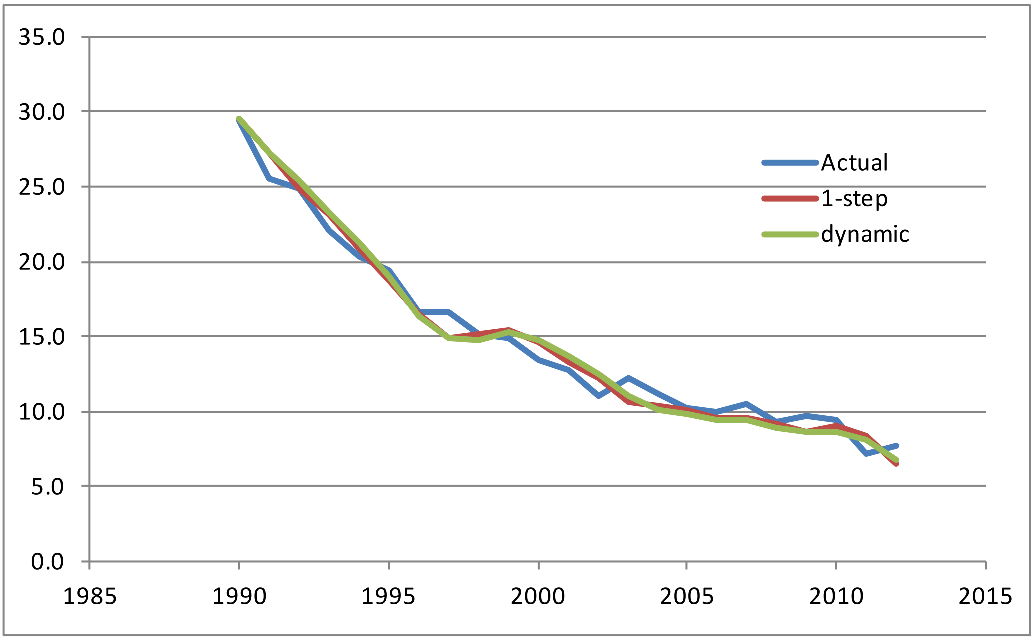
Figure 13 shows the actual and estimated values of the dependent variable F/T.

The model's dynamic estimate for total road deaths between 1990 and 2012 is 10,919, compared to the actual figure of 11,004 and 11,014 in Model 1. The sum of absolute errors from the dynamic process is 686, an average of 30 fatalities per annum, which is marginally worse than under Model 1.

The last column in Table 5 is analogous to that in Table 3, showing how each variable contributed to the decline in road fatalities between 1990 and 2012. The results strongly reinforce those obtained in Model 1 and have very similar standard errors. About 47%

of the decline in fatalities can be attributed to improvements in vehicle quality (treating motorcycles as less safe vehicles). Road quality contributed 18% and driver-related factors contributed 35%.

Figure 13: Model 2 Results for F/T (number per billion km)



The short term effect of changes in TARPs accounts for only about 3% of the total TARPs effect and it is of questionable significance. However, there is an additional small, but statistically significant contribution from annual changes in CBT which account for 93 fewer deaths over the period.[19] As in Model 1 though, the dummy variable for the undercount of CBT and MBT in 2006/07 is not significant. Thus we infer that the main mechanism through which breath testing affects fatalities is through a gradual process of changing driver behaviour. Annual variation in the amount of breath testing is of relatively little consequence (as was quarterly variation in KSSCW).

Table 6 compares the coefficients from the two models in terms of their effect on the number of fatalities, evaluated at mean travel over the period 1990-2012, for a change in the independent variables of one standard deviation.

The marginal effects of road investment and motorcycle penetration are closer than their actual effects over 1990-2012 (refer Tables 3 and 5), suggesting that the former has more potential to reduce fatalities than it has had historically. The same is true for vehicle crash worthiness.

In contrast, most of the historical effect of breath testing on fatalities has come from the sheer number of tests undertaken, rather than from a high unit effect. Similarly for advertising (as measured by TARPs).

Mean rural speed is another variable that would seem to have more potential to reduce the number of fatalities, although the trade-off with travel time would need to be considered.

[19] Note that it may be inconsistent to be holding CBT changes at zero while TBT is allowed to vary.

Table 6: Effect Sizes
(Effect on Fatalities for Change in Variable of One Standard Deviation)

	Model 1	Model 2
Road investment/travel	-24.0	-26.4
Vehicle crash worthiness	16.9	18.5
Motorcycle registrations/travel	28.7	33.6
Total breath tests (CBT+MBT)	-8.1	-7.8
TARPs	-7.9	-8.3
Mean rural speed	15.0	15.8

3.3 Caveats

No evaluation is perfect. As discussed above, many variables have been spliced and estimated in various ways, so measurement error presents a source of bias. For example, quite apart from the error in measuring real (inflation adjusted) investment in roading infrastructure, it is only valid as a proxy for road quality if it has a reasonable correlation with expenditure on safety. Similarly, year of manufacture is not a perfect measure of vehicle crash worthiness.

Although the models provide an explanation for most of the decline in the F/T trend, there are other factors that could also affect F/T that we have not been able to include.

Medical care: The promptness and quality of medical care for accident victims could be important. While it will not affect the probability of a serious accident, it may affect the probability of death, given a serious accident. Sanchez-Magnas et al (2009)[20] find that a reduction of ten minutes in the time between an accident and the arrival of a medical response has a statistically significant association, with an average decrease of one third in the probability of death. There is no doubt that there has been an improvement in response times over the past 20 years, partly due to improved mobile phone access and coverage.

Vehicle occupancy: Changes in vehicle occupancy could also affect F/T, as the probability of a fatality, given an accident, may be higher if there are more passengers, as would be the probability of a reported injury crash. Unfortunately time series data on vehicle occupancy is fragmented, although there is little indication of change from Travel Survey data.

Type of travel: Changes in the travel mix could affect the rate of fatalities even if the total amount of travel does not change. For example in KSSCW it is noted that around 15% of road deaths occur in crashes involving trucks even though trucks constitute only 6% of distance travelled. Thus a relative change in truck travel (vkm) could have a disproportionate effect on fatalities. The rural-urban mix of travel could also be important, although this will be partly affected by expenditure on road construction.

Interaction effects: We have not considered interaction effects as degrees of freedom are rather precious in the above models, but limited testing might be worth pursuing in cases where there is a strong theoretical prior for interactive effects. For example the Supplementary Road Safety Package in the 1990s (see Guria and Leung, 2004) relied on combining more enforcement with graphic advertising.

[20] Sanchez-Magnas, R., Garcia-Ferrer, A., de Juan, A., Arroyo, A.M. (2009): The probability of death in road traffic accidents. How important is a quick medical response? *Accident Analysis and Prevention*, 42, 1048-1056.

Another point to recognise is that the 'explanatory' variables merely provide one level of explanation. They are in turn explained by other variables. The rate at which safer vehicles enter the fleet is related to the rate of growth in per capita income and to regulations. And while it is clear that a higher proportion of motorcycles in the vehicle fleet is likely to raise the number of road fatalities, this tells us nothing about who or what causes the crashes or why people buy motorcycles. Understanding why the explanatory variables changed as they did between 1990 and 2012 is altogether another research topic.

Finally the above models are premised on weak simultaneity. A solid body of theory suggests that correlation between our explanatory variables and F/T indicates a causative effect from the explanatory variables to F/T, but it may be the case that a change in F/T leads or prompts a change in the explanatory variables. In a general sense this will always be true – if there were no more road deaths from now on we could expect the rate of improvement in vehicle crash worthiness to abate and roading expenditure would be re-prioritised. Enforcement and advertising would decline.

Econometrically though we need only require that annual changes in F/T do not prompt sudden marked changes in the explanatory variables. In 2010 for example, the police began enforcing a 4 km/h speed tolerance during holiday weekends, in response to a high Queen's Birthday weekend road toll in that year. This policy was found in KSSCW to have no significant effect (to date), and in any case it is not a variable in the above models. One might imagine though, that enforcement and/or advertising blitzes could have occurred in the past after a period with a high number of road deaths. While it is unlikely that such endogeneity would operate on annual time scale, it may nonetheless warrant further investigation.

3.4 Refinements

Based on the points above there are some issues that could be further researched:

1. Interaction effects, such as between advertising (TARPs) and breath testing.
2. The effect of the ban on the use of cell phones while driving, introduced in 2009, although one would first want to see some evidence of a measurable effect, such as whether there has been a decline in the number of accidents for which cell phone use was a major contributing factor.
3. Analysis of possible endogeneity between changes in road safety policies and previous changes in the rate of fatalities.
4. The effects of changes in the mix of travel: urban-rural, age of driver, heavy versus light vehicles, congested versus uncongested travel, and so on.
5. The above models investigate F/T. Complementary research could perhaps look at travel per vehicle (T/V) and at vehicles per capita (V/P).

Appendix A: Lagged Effects

To circumvent the problem of multicollinearity and preserve degrees of freedom when the right hand side of an equation needs to incorporate lagged effects, the lag structure is frequently approximated by some polynomial function. The Almon distributed lag structure is given by:

$$Y_t = \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \mu_t$$

- where the β may be linked via a polynomial function.

In particular, Koyck extended the idea of linking the β by postulating a geometrically declining lag function, giving rise to an equation of the form:

$$Y_t = \beta(w_0 + w_1 L + w_2 L^2 + \ldots)X_t + \mu_t$$

- where L is the lag operator such that $L^i X_t$ denotes $X_{t-i.}$ and where $w_i = (1-\lambda)\lambda^i$.

The equation then condenses to:

$$Y_t = \beta(1-\lambda)X_t + \lambda Y_{t-1} + \mu_t - \lambda\mu_{t-1} \qquad (2)$$

However, a lagged dependent variable in an equation similar to (2) would also arise in a situation of *partial adjustment* or *adaptive expectations*. Under the former for example, motorists may take some time to change their behaviour in response to new road safety measures.

$$Y_t^* = \alpha + \beta X_t$$

- where Y* represents the optimal value of Y in response to given X. Y* cannot be observed, but a relationship between it and the observed Y might be given by:

$$Y_t - Y_{t-1} = \gamma(Y_t^* - Y_{t-1}) + \mu_t \qquad 0 < \gamma <= 1$$

Amalgamating these equations yields:

$$Y_t = \alpha\gamma + \beta\gamma X_t + (1-\gamma)Y_{t-1} + \mu_t \qquad (3)$$

Under the *adaptive expectations* model motorists adjust their behaviour according to their expectations about road safety. Expectations X* of the exogenous variable X are based on past observation. Thus:

$$X_t^* - X_{t-1}^* = \delta(X_t - X^*_{t-1}) \qquad 0 < \delta <= 1$$

Given also $Y_t = \alpha + \beta X^*_t + \mu_t$

With some algebraic manipulation these can be amalgamated to yield:

$$Y_t = \alpha\delta + \beta\delta X_t + (1-\delta)Y_{t-1} + \mu_t - (1-\delta)\mu_{t-1} \qquad (4)$$

Equations (2) - (4) all have the same form, differing only in their error structure, although for estimation purposes this difference is not trivial.

The three situations given by equations are not mutually exclusive. Indeed they are not even exhaustive of the possibilities. Drivers' behaviour may be based on expectations about road safety measures (not just on what is actually implemented), their behaviour may be slow to respond to new measures, or the measures may have a compounding (or memory) effect - they need to be seen as more than transitory before they induce a behavioural response.

With regard to equation (3) for example, the effect of X on Y can be determined by estimating σ and θ in the equation below and calculating the desired β from equation (3).

$$Y_t = \omega + \sigma Y_{t-1} + \theta X_t + \mu_t$$

For this equation OLS yields consistent and efficient estimates for uncorrelated errors. However the estimates will be biased, although correcting for bias often leads to worse estimates overall (such as in terms of mean squared error) due to an accompanying increase in variance.[21]

Thus this equation, or with the addition of serially correlated errors, is our preferred specification for a long term model of the trend in F/T.

[21] Ignoring orders of magnitude smaller than n^{-2}, the bias is approximately equal to $-2/n$.