Denne artikel er publiceret i det elektroniske tidsskrift
Artikler fra Trafikdage på Aalborg Universitet
(Proceedings from the Annual Transport Conference
at Aalborg University)
ISSN 1603-9696
www.trafikdage.dk/artikelarkiv

Why is it so difficult to explain the decline in traffic fatalities?

Rune Elvik (re@toi.no)
Transportøkonomisk institutt

Abstrakt

In many highly motorised countries, the number of traffic fatalities has gone down by about 80 percent since the peak number, which was reached around 1970. What explains this decline? Is it principally the result of road safety policy, or have other factors made a larger contribution? This paper argues that it is difficult to give a scientifically rigorous explanation of the decline in traffic fatalities. There are five main problems: (1) There are very many potentially relevant explanatory variables. (2) Some of the relevant explanatory variables change slowly at an almost constant rate. (3) Data are incomplete or missing about many potentially relevant variables. (4) Some variables are affected by measurement errors or discontinuities in time series. (5) Many of the explanatory variables are very highly correlated with each other and with time. These problems are illustrated using Norway as an example. It is shown that the problems listed above can result in models that are non-sensical although they pass formal tests of model quality. The lesson is that one should never judge how good a model is merely in terms of formal criteria. Some strategies for developing more meaningful models are discussed.

1 Introduction

The number of traffic fatalities reached an all-time high in many highly motorised countries around 1970. Since then, the number of traffic fatalities has declined substantially in many countries. As an example, Figure 1 shows the number of traffic fatalities in Norway from 1970 to 2015. The number recorded in 1970, 560, was the highest ever. The number recorded in 2015, 117, is the lowest since 1947. The decline from 1970 to 2015 was 79 percent.

Similar reductions have occurred in many highly motorised countries. In Sweden, traffic fatalities declined from 1307 in 1970 to 259 in 2015 (80 percent). In Denmark, traffic fatalities declined from 1213 in 1971 to 167 in 2012 (86 percent). In Great Britain, traffic fatalities declined from 7763 in 1972 to 1713 in 2013 (78 percent). The overall trend has been similar in many other countries.

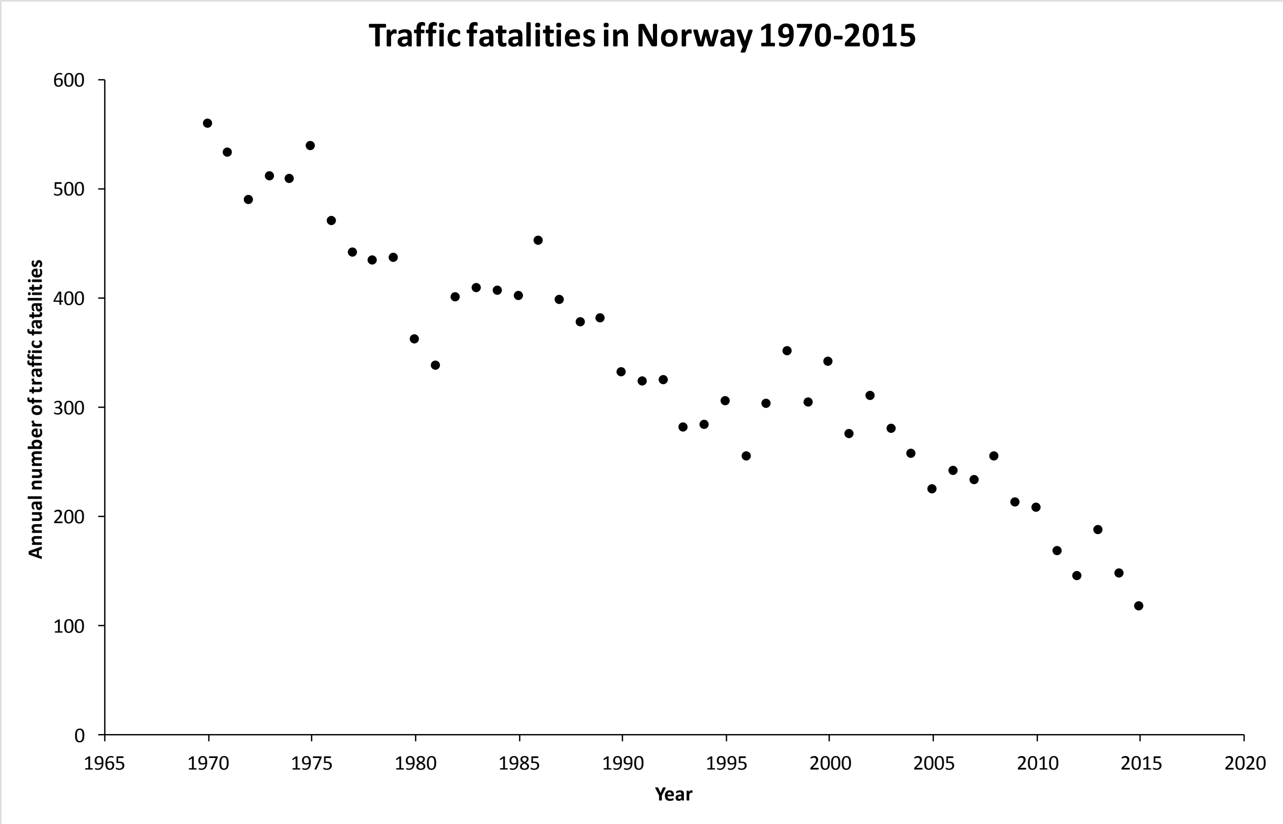

Figure 1: Traffic fatalities in Norway 1970-2015

Globally, many countries are still in an early phase of motorisation. In these countries, the number of traffic fatalities can be expected to grow if they follow the same historical development as the currently highly motorised countries. If, by contrast, the factors that have contributed to the decline in traffic fatalities in many highly motorised countries can be identified, countries that are still early in motorisation may perhaps benefit from this knowledge and avoid, or at least reduce, the increase in traffic fatalities that occurred until about 1970 in many highly motorised countries.

Unfortunately, explaining the decline in traffic fatalities in the highly motorised countries is surprisingly difficult. The objective of this paper is twofold. First, to point out some reasons why it is difficult to explain the decline in traffic fatalities. Second, to show by means of simple examples that paying insufficient attention to the problems may result in models that make little sense, although the models are formally good in terms of criteria such as goodness-of-fit, normality of residuals, homoscedasticity of residuals, and so on.

2 Identifying potentially relevant explanatory variables

The first task in trying to explain the decline in traffic fatalities is to identify potentially relevant explanatory variables. This is no small task. Very many variables influence road safety and the number of traffic fatalities.

If one relies on annual data, the maximum number of observations attainable in a study seeking to explain the decline in traffic fatalities in a given country is about 45-50. In a model based on 45-50 observations one cannot hope to reliably estimate the effects of more than, say, 5-10 explanatory variables. However, the number of variables influencing traffic fatalities is considerably larger. Table 1 lists variables that have been found to be related to traffic fatalities in Norway, based on studies relying on Norwegian data.

A total of 31 variables are listed in Table 1. The table is obviously not complete. If one assumes that factors affecting traffic fatalities have similar effects in all countries, studies from all over the world become relevant. The list in Table 1 could then be expanded to, literally, several hundred variables. There is no chance of estimating the statistical relationship between all these variables and the number of traffic fatalities. Any model containing a selection of variables entails an unknown, but potentially great, risk of omitted variable bias.

3 Variables changing at a constant rate over time

Some of the variables influencing the number of fatalities change at a slow and fairly constant rate. These variables may not vary enough from year to year for their effect to be reliably estimated. Due to the very strong correlation between slowly changing variables and time, inclusion of such variables in a model also containing year as an explanatory variable is problematic. As an example of a variable showing a fairly stable development over time, consider Figure 2.

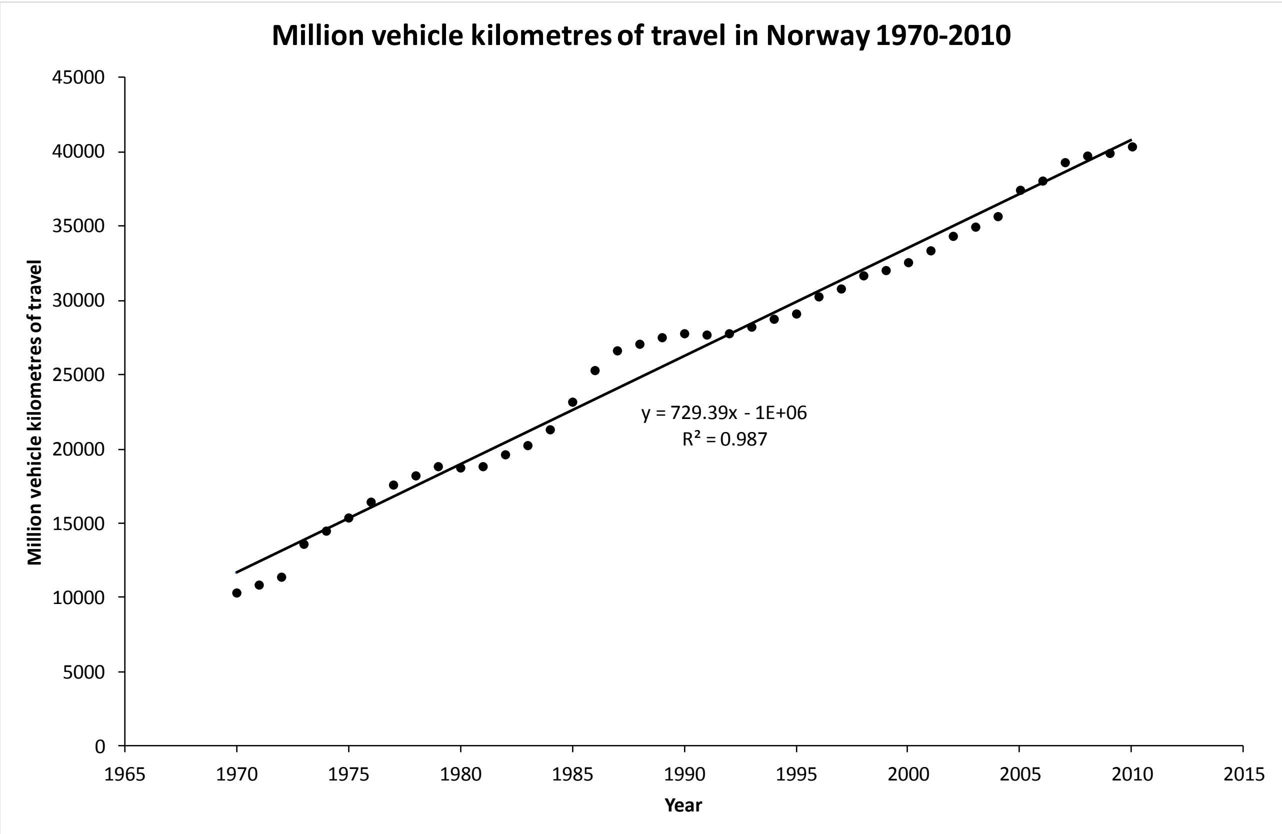

Figure 2: Vehicle kilometres of travel in Norway 1970-2010

Table 1: Factors that have been found to influence the number of traffic fatalities in Norway

Main category of variables	Variables found to be related to traffic fatalities (number)	Potential relevance to explanation of decline in fatalities	Studies finding relationship
Daylight	(1) Minutes of daylight	Changed as daylight savings time was introduced in 1980, and extended from end of September to end of October in 1996	Fridstrøm et al. 1995
Weather	(2) Monthly days with snowfall	May change gradually over time as a result of global warming	Fridstrøm et al. 1995
	(3) Snow-depth	May change gradually over time as a result of global warming; may lose some of its protective effect due to better protective systems in cars	Fridstrøm et al. 1995; Elvik 2016A
	(4) Monthly days with rainfall	Rain has become more frequent over time; this by itself may change its relationship to traffic fatalities	Fridstrøm et al. 1995; Elvik 2016A
Real income	(5) Income per capita, fixed prices	Rising income is strongly associated with increased travel	Elvik 2015A
Unemployment	(6) Unemployment as percentage of labour force	Increasing unemployment is associated with a decline in traffic fatalities; unemployment is a short-term factor that may explain fluctuations around the long-term trend	Fridstrøm 1999, Elvik 2015A
Population density	(7) Inhabitants per square km	Population density increases over time as population grows; urbanisation means that a larger share of the population lives in densely populated areas	Fridstrøm 1999
Women pregnant	(8) Women pregnant in first quarter per 1,000 women	There have been large changes over time in pregnancies, as the reproduction rate has tended to go down after the "baby-boom" between 1945 and 1970	Fridstrøm 1999
Urban planning	(9) Design of street network in residential areas	The principles guiding the design of the street network in urban areas have changed over time; safety is related to these principles	Muskaug 1980
Exposure	(10) Total vehicle kilometres driven	There is a positive relationship between vehicle kilometres and fatalities	Høye 2014
	(11) Heavy vehicle share	The share of vehicle kilometres performed by heavy vehicles	Langeland and Phillips 2016
	(12) Pedestrian and cyclist kilometres of travel	Pedestrian and cyclist kilometres of travel have fluctuated over time, in particular cyclist travel	Elvik 2005
	(13) Novice driver kilometres of travel	The share of driving performed by novice drivers has changed over time, reaching a peak in the 18980s, declining after that	Elvik 2005
	(14) Traffic density	Vehicle kilometres per kilometre of road; has tended to increase over time	Fridstrøm 1999

Table 1: Factors that have been found to influence the number of traffic fatalities in Norway

Main category of variables	Variables found to be related to traffic fatalities (number)	Potential relevance to explanation of decline in fatalities	Studies finding relationship
	(15) Bus transport supply	Bus kilometres per kilometre of road; has tended to increase over time	Fridstrøm 1999
	(16) Mean age of cars	New cars have more safety features than old cars; the turnover rate for cars determines how quickly these features reach full penetration	Fridstrøm 1999
	(17) Share of cars with electronic stability control	Electronic stability control reduces the number of loss-of-control accidents and the share of cars having the system has increased rapidly after about 1995	Høye 2011; Elvik 2015B
Road safety measures	(18) Adoption of a quantified road safety target	Adopting a quantified road safety target is associated with an accelerated decline in traffic fatalities	Allsop et al. 2011
	(19) Changes in speed limits	There were major changes in speed limits during 1978-80 and in 2001	Sakshaug 1986; Ragnøy 2004
	(20) Law on seat belt wearing	Introduced without fines in 1975, with fines in 1979	Fridstrøm 1999
	(21) Speed cameras and section control	Speed cameras were introduced in 1988; section control in 2004; both systems have been extended in recent years	Høye 2015A, 2015B
	(22) Converting junctions to roundabouts	Converting junctions to roundabouts reduces accident severity; more than 1000 junctions have been converted to roundabouts	Odberg 1996; Tran 1999
	(23) Building motorways	Share of all vehicle kilometres driven on motorways	Elvik 2005
	(24) Introducing road lighting	Road lighting reduces the number of fatal accidents in darkness; it has been extended after 1970	Wanvik 2009
	(25) Level of enforcement	Fixed penalties issued per million vehicle kilometres; has tended to decline over time	Elvik 2005
	(26) Level of fixed penalties	Increases in fixed penalties may deter traffic violations; many violations are associated with increased fatality rate	Elvik 2016B
Road user behaviour	(27) Mean speed of traffic	Changes in the mean speed of traffic are associated with changes in the number of fatalities; speed tended to increase until 2006, thereafter decline	Elvik 2009, 2013
	(28) Drinking and driving	Road side surveys made in selected years indicate a decline in drinking and driving and the share of fatalities involving drinking drivers	Christophersen et al. 2016

Table 1: Factors that have been found to influence the number of traffic fatalities in Norway

Main category of variables	Variables found to be related to traffic fatalities (number)	Potential relevance to explanation of decline in fatalities	Studies finding relationship
	(29) Seat belt wearing	Increased seat belt wearing is associated with fewer traffic fatalities	Fridstrøm 1999; Høye 2016
	(30) Use of child restraints	The use of child restraints has increased over time; child restraints reduce the risk of fatal injury to children in accidents	Høye et al. 2015
	(31) Driving under the influence of drugs	Road side surveys indicate changes in driving under the influence of drugs as well as the fatality risk associated with such driving	Gjerde et al. 2011, 2013

Figure 2 shows the development of vehicle kilometres of travel in Norway from 1970 to 2010. Although the annual changes vary a little, their correlation with time is almost perfect. A negative binomial regression model was run, using the natural logarithm of vehicle kilometres as independent variable. A coefficient for ln(vehiclekm) of –0.622 was estimated with a standard error of 0.0230. A model using year as the only independent variable was then run. The coefficient for year was –0.021 (standard error = 0.0008). Finally, a model including both variables was run. The coefficient for year (standard error in parentheses) was –0.022 (0.0029). The coefficient for ln(vehiclekm) was 0.028 (0.0876).

Including both variables in the same model clearly makes no sense. However, omission of one of the variables is almost bound to generate omitted variable bias, since the omitted variable is correlated both with one or more of the independent variables included in the model and the dependent variable.

In the simple example given here, it was possible to examine what happened when a variable was included or excluded from a model. A considerably more serious uncertainty is introduced by having to omit variables with incomplete or missing data.

4 Variables with incomplete or missing data

Data are missing or incomplete for very many variables that influence the number of traffic fatalities, including some variables that are likely to be important. It is easy to give examples of such variables.

In 1970, Norway was still in a comparatively early phase of motorisation, compared to countries like Sweden or the United States. It is therefore likely that an average driver in Norway in 1970 was less experienced than an average driver is today. Drivers who started their driving careers early may now benefit from 40-50 years of experience, which very few drivers had in 1970. The mean driving experience of the population of drivers has probably grown steadily from 1970 until now, but it is impossible to reconstruct the historical development of this potentially important variable. Moreover, even if a historical reconstruction were possible, the variable would likely be almost perfectly correlated with time, creating the same problems of estimation as shown above for vehicle kilometres of travel.

It is widely agreed that road user behaviour is important for safety. Data on the development over time of road user behaviour is incomplete. In Norway, seat belt wearing among car drivers has been monitored since 1973. Data are missing for 1970-1972, 1989, 1992, 1994 and 1996. Monitoring of seat belt wearing among rear seat passengers was discontinued in 2005.

Comparable data on the mean speed of traffic exist only from 2006 onwards. For years before 2006, data are only sporadically available and any reconstruction based on these data will be incomplete (Elvik 2012). The same goes for drinking and driving (Christophersen et al. 2016). Roadside surveys were made in 1971, 1977, 1981-82, 2005-06 and 2008-09. However, these surveys were quite different and are strictly speaking not comparable, except perhaps for the two most recent surveys. One may consider using a proxy variable to indicate drinking and driving, but the variables that are available as proxies are likely to be misleading.

Seat belt wearing, speed and drinking and driving are just three of very many variables describing road user behaviour. It is therefore clear that missing data about potentially important variables is a big problem when trying to explain the decline in traffic fatalities.

5 Discontinuities in time-series and errors in variables

Some of the time series that are available for all years after 1970 in Norway have discontinuities. Thus, estimated vehicle kilometres of travel has a discontinuity in 1997. Motorway length has a discontinuity in 2005. Presumably, the data available for the years after these discontinuities are of better quality than the older data. However, that means that by relying on older data for the years before the discontinuities, one will be using data that are known to have measurement errors.

While there have been few roadside surveys of drinking and driving in Norway, there is another source of data that might indicate changes over time in drinking and driving. The Traffic Police records how many drinking drivers they detected as a result of enforcement. A time-series showing the number of drinking drivers detected by the police per 1,000 drivers who were checked can be created from 1980 onwards. Figure 3 shows this time series.

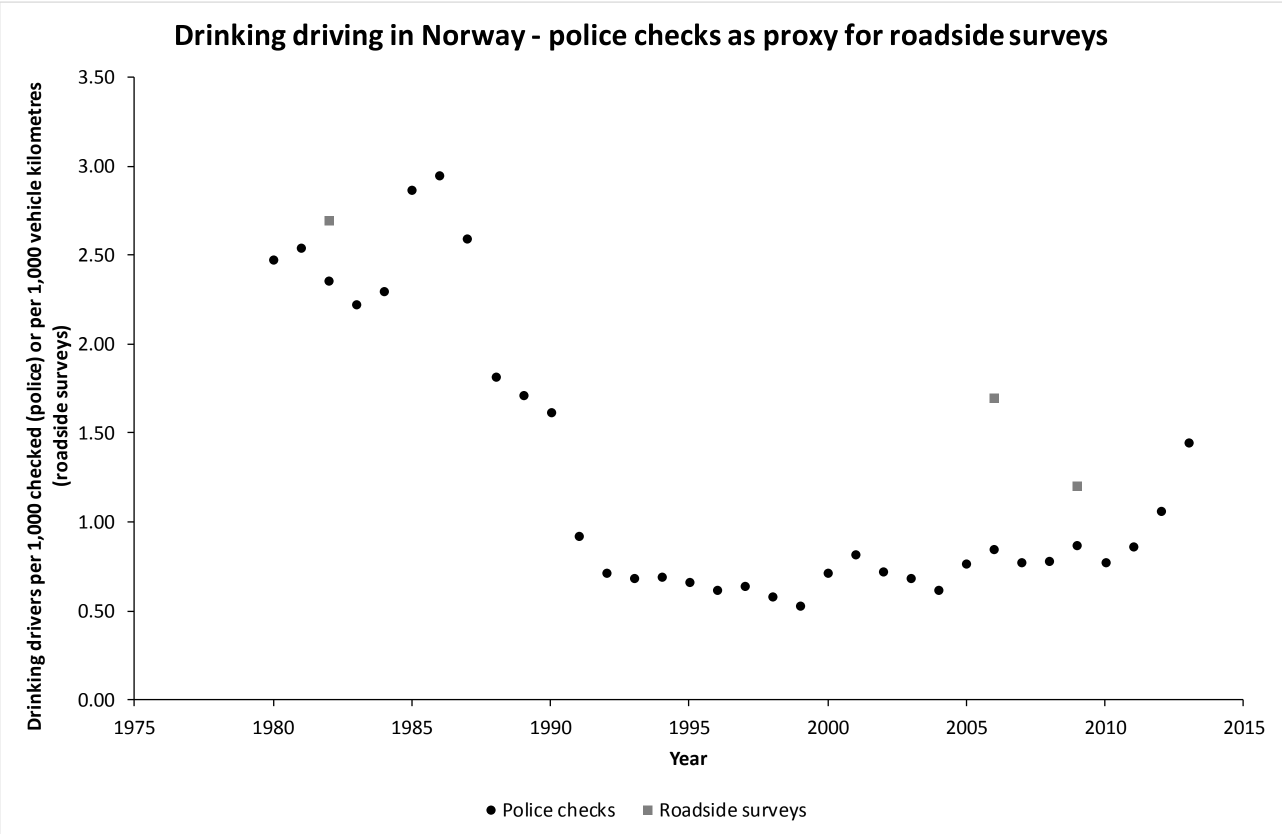

Figure 3: Drinking and driving: Police data as proxy for roadside surveys

Figure 3 also shows the estimated share of vehicle kilometres driven by drivers having a blood alcohol concentration of 0.05 percent or more in three of the roadside surveys. While the police data for 1982 are fairly close to the estimate based on the roadside survey, the estimates for 2006 and 2009 are far apart and show inconsistent changes (reduction in drinking and driving according to the roadside surveys; a barely perceptible increase according to police data).

Police data indicate a large increase in drinking and driving in recent years. This is probably an artefact. It is only recently that the police started to do routine breath testing of all drivers they stop. Earlier, drivers were only tested if the police suspected them of drinking and driving. Thus, the police data are not comparable over time and the “gold standard” of roadside surveys provides too few data points to give a basis for judging the accuracy of police data.

6 Correlations between variables

Data have been collected on 15 variables that may influence the number of traffic fatalities for the period 1997-2013. The year 1997 was chosen as first year in order to avoid the discontinuity in the time-series for vehicle kilometres of travel, mentioned in section 5 of the paper. Table 2 lists these variables.

Table 3 shows the correlation between the variables. The table contains a total of 120 correlation coefficients for pairs of variables. Nearly half of these, 59, indicate stronger correlations than 0.7 or –0.7. 25 of the correlations are stronger than 0.9 or –0.9. These correlations are close to perfect co-linearity and may present problems when the correlated variables are included in a model intended to explains changes in the number of traffic fatalities.

This raises the issue of how best to select variables for inclusion in a model intended to explain the decline in the number of traffic fatalities. Not everybody agrees that co-linearity between explanatory variables is a problem. Thus, Fridstrøm explains (2015, page 17): “Non-experimental data are notoriously interrelated or at least correlated, i.e. collinear. In fact, collinearity is the very reason why we need multiple regression analysis to understand what is going on. It makes no sense at all to require that collinearity be avoided. (However), when several relevant variables are collinear, it is hard to estimate their respective partial effects. The estimates will be imprecise. But this will be reflected in the estimated standard errors, the t-tests, the p-values, and so on.” Large standard errors may thus indicate co-linearity. The next section presents two possible solutions to the problem and shows that the resulting models make little sense.

7 Models that make no sense satisfy formal criteria of goodness

One approach that can be taken to the problem of correlations between explanatory variables, is to develop a model by selecting variables that are moderately correlated with each other as explanatory variables. Based on the correlations in Table 3, such a model was developed containing the following independent variables: (1) Share of vehicle kilometres performed by heavy vehicles, (2) Unemployment as percentage of labour force, (3) Share of 18 year olds obtaining a driving licence, (4) Drivers cited for traffic offences per million vehicle kilometres of travel, (5) Drivers testing positive for alcohol per 1,000 drivers checked by the Traffic Police, (6) Precipitation as percentage of normal annual amount. The left panel of Table 4 shows the coefficients that were estimated.

Is this a good model? To answer this question, the following criteria of model quality have been applied:

Table 2: Potential explantory variables for decline in traffic fatalities

Abbreviated name	Full name and description
Yrcount	Count of years from 1997 (= 1) to 2013 (= 17)
Millkm	Million vehicle kilometres of travel
Heavyshare	Percentage of all vehicle kilometres performed by heavy vehicles (> 3.5 metric tons)
Mcshare	Percentage of all vehicle kilometres performed by motorcycles
Beltuse	Percentage of car drivers wearing seat belts
ShareESC	Percentage of all car kilometres driven by cars with electronic stability control
Sharefive	Percentage of all car kilometres driven by cars with five stars according to EuroNCAP
Sharebrake	Percentage of all car kilometres driven by cars with emergency braking system
Unemploy	Unemployment as percentage of the labour force
Youngdrive	Percentage of 18 year olds who obtain a driving licence at the age of 18
Checkmill	Drivers checked by the police per million vehicle kilometres of travel
Ticketmill	Number of drivers cited for traffic offences per million vehicle kilometres of travel.
UPdrunk	Drivers testing positive for alcohol per 1,000 drivers checked by the Traffic Police
Kmmedian	Kilometres of road with median guard rail
Precip	Annual precipitation as percentage of normal amount

Table 3: Correlations between variables

	Bivariate correlations Pearson's r														
	Yrcount	Millkm	Heavyshare	Mcshare	Beltuse	ShareESC	Sharefive	Sharebrake	Unemploy	Youngdrive	Checkmill	Ticketmill	UPdrunk	Kmmedian	Precip
Fatals	-0.922	-0.910	0.414	-0.896	-0.791	-0.924	-0.892	-0.922	0.014	0.643	0.706	-0.561	-0.683	-0.915	-0.154
Yrcount	1	0.995	-0.470	0.967	0.882	0.995	0.950	0.988	-0.160	-0.781	-0.861	0.562	0.759	0.959	0.051
Millkm		1	-0.493	0.960	0.864	0.986	0.926	0.973	-0.179	-0.816	-0.863	0.607	0.724	0.937	0.055
Heavyshare			1	-0.512	-0.451	-0.422	-0.292	-0.387	-0.015	0.553	0.535	-0.173	-0.316	-0.290	0.116
McShare				1	0.857	0.947	0.886	0.937	-0.122	-0.731	-0.917	0.440	0.701	0.892	0.043
Beltuse					1	0.879	0.861	0.882	-0.246	-0.677	-0.793	0.377	0.784	0.857	-0.005
ShareESC						1	0.975	0.998	-0.193	-0.775	-0.829	0.557	0.765	0.979	0.051
Sharefive							1	0.986	-0.256	-0.706	-0.753	0.475	0.785	0.993	0.055
Sharebrake								1	-0.192	-0.743	-0.814	0.526	0.788	0.990	0.050
Unemploy									1	0.542	0.216	-0.101	-0.100	-0.202	-0.083
Youngdrive										1	0.715	-0.675	-0.419	-0.684	-0.014
Checkmill											1	-0.320	-0.689	0.509	0.170
Ticketmill												1	0.266	0.509	0.170
UPdrunk													1	0.808	0.087
Kmmedian														1	0.104
Precip															1

Table 4: Estimated models – coefficients, standard errors and P-values

	Panel A: Model 1			Panel B: Model 2		
Variables and terms	**Estimate**	**Standard error**	**P-value**	**Estimate**	**Standard error**	**P-value**
Constant term	4.795	0.529	0.000	8.255	0.735	0.000
Share of heavy vehicles	-0.192	0.075	0.010	0.230	0.095	0.016
Unemployment	-0.246	0.058	0.000	0.155	0.079	0.051
Young drivers at 18	0.056	0.013	0.000	-0.066	0.022	0.002
Drivers cited per million vehicle km	0.060	0.043	0.163	-0.113	0.045	0.011
Drivers testing positive for alcohol	-0.560	0.107	0.000	0.361	0.181	0.046
Precipitation as percent of normal	-0.003	0.002	0.067			
Share of cars with electronic stability control				-0.017	0.003	0.000
Over-dispersion parameter	0.001			0.000		
Percent of systematic variation explained	83.1			99.2		

Table 5: Assessing models in terms of criteria of model quality

Criteria for evaluating model quality	Assessment for model 1	Assessment for model 2
Unbiased model prediction	4293 fatalities predicted, 4295 recorded: model is unbiased	4296 fatalities predicted, 4295 recorded: model is unbiased
Direction of annual changes correctly modelled	10 correct, 6 incorrect: model not satisfactory	11 correct, 5 incorrect: model not satisfactory
Statistical significance of regression coefficients	5 of 7 significant at P = 0.05: quite satisfactory	6 of 7 significant at P = 0.05: satisfactory
Share of systematic variation explained by model	83.1 % explained: model can probably be improved	99.2 % explained: nearly all systematic variation explained
Normality of standardised residuals	Chi-square test shows significant deviation from normality	Chi-square test indicates no deviation from normality
Homoscedasticity of residual terms	Difference in slopes 0.016; standard error 0.011: homoscedastic	Difference in slopes 0.005; standard error 0.006: homoscedastic
Autocorrelation of residual terms	No significant autocorrelation for lags 1 through 15	Significant autocorrelation for lag 1, not for lags 2 through 15

1. The model should make unbiased prediction, i.e. it should not predict too many or too few fatalities.
2. The model should track annual changes in fatalities, i.e. predict a decline when there was one and predict an increase when there was one.
3. Estimated regression coefficients should be precise and preferably statistically significant; large standard error may indicate co-linearity.
4. The model should explain as much as possible of the systematic variation in the number of fatalities, but not be over-fitted, i.e. explain part of the random variation in the number of fatalities in addition to the systematic.
5. The standardised residual terms should have a normal distribution.
6. The standardised residual terms should be homoscedastic.
7. The residual terms should not be autocorrelated.

Table 5 assesses the model in terms of these criteria. Although unbiased, the model does not track annual changes in the number of fatalities very well. Figure 4 shows the recorded number of fatalities and model estimates year-by-year.

Figure 4: Traffic fatalities in Norway 1997-2013 and model to explain annual changes

The model is consistent with the direction of annual changes in 10 cases, inconsistent in 6 cases. The estimated regression coefficients are quite precise; 5 of 7 are statistically significant the 5 % level, a sixth coefficient is statistically significant at the 10 % level. If large standard errors indicate co-linearity, these results are reassuring. The model explains 83 % of the systematic variation in the number of fatalities; in principle it ought to be possible to improve this. The standardised residuals are not normally distributed. They are, however, homoscedastic. A simple test of homoscedasticity, suitable for graphical representation, was used to determine this. The logic of the test can be explained by reference to Figure 5.

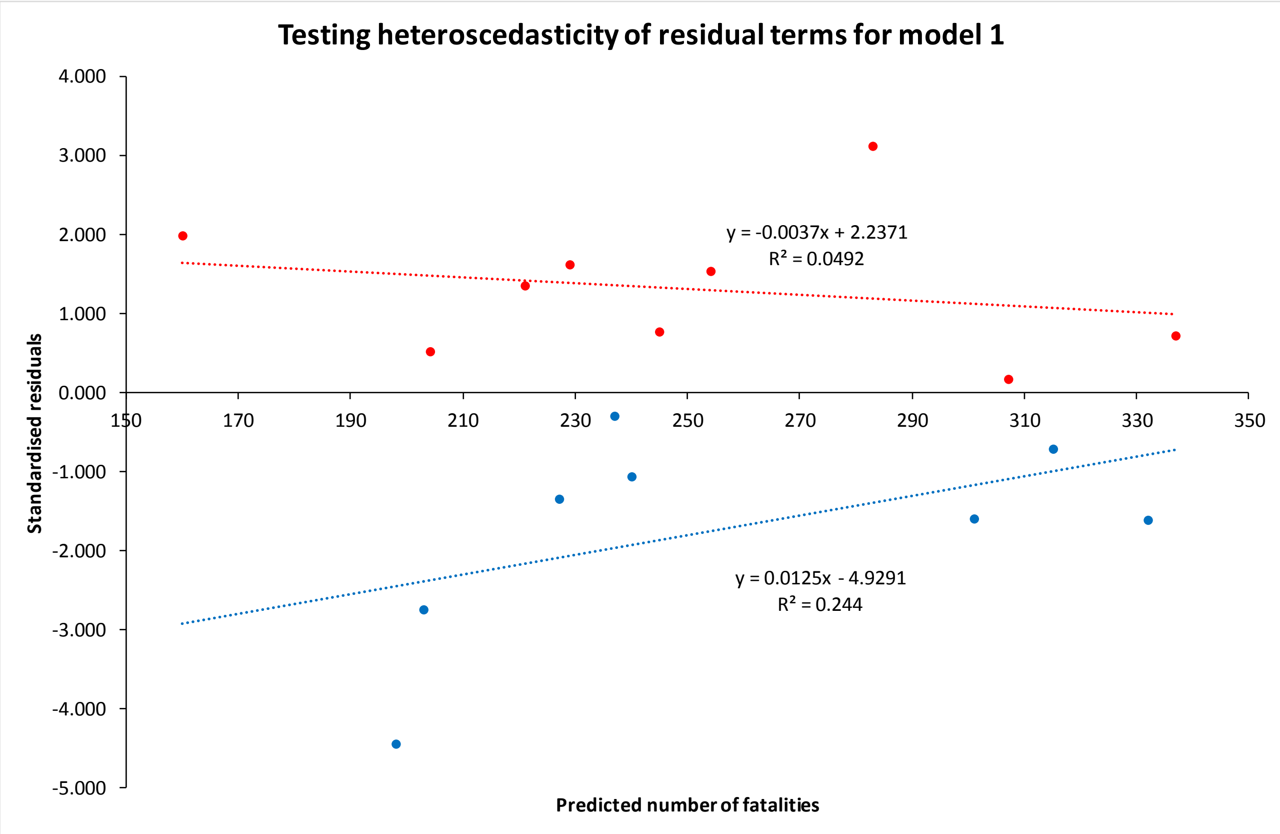

Figure 5: Testing heteroscedasticity of residual terms

Positive residuals are shown in the upper half of the Figure, negative residuals in the lower half. Trend lines have been fitted to the residuals. Ideally speaking, if residuals are perfectly homoscedastic, these lines should be horizontal. In Figure 5, both lines have a slope, but in opposite directions. This indicates heteroscedasticity. The slope coefficients have standard errors (not shown in Figure 5); these were applied to test whether the difference in slopes was statistically significant. For Figure 5, the difference in slopes was 0.0162 (0.0125 – (-0.0037)). The standard error of this difference was 0.0109, which suggests that there was no significant difference between the slopes and therefore no significant heteroscedasticity. Finally, as far as autocorrelation of the residuals is concerned, no significant autocorrelation was found.

On the whole, therefore, the model is, if not perfect, at least satisfactory. It gives unbiased predictions; most of the coefficients are statistically significant; the residuals are homoscedastic and have no significant autocorrelation. On the other hand, one would like a good model to contain variables that are believed to be important in influencing the number of fatalities, not just variables that are comparatively uncorrelated. Developing regression models in road safety is very much a process of trial and error (Hauer 2015), a fact one should never try to disguise. Road safety is such a complex phenomenon, that one cannot hope to develop good explanatory models without relying on extensive exploratory analysis.

To see if a better model could be developed, variables were therefore added to model 1 one-by-one until the model became over-fitted. The precipitation variable was mostly not significant and was therefore dropped. Model 2, also presented in Table 4, explained 99.2 % of the systematic variation in the number of traffic fatalities in Norway between 1997 and 2013. The only difference between models 1 and 2, is that in model 2 the precipitation variable has been replaced by a variable showing the share of cars having electronic stability control. Model 2 is better than model 1 according to nearly all the criteria of model quality; the only one where it scores marginally worse than model 1 is autocorrelation of residual terms.

Does this mean that model 2 should be trusted and model 1 rejected? For the variables the two models have in common, all coefficients in model 2 have the opposite sign of model 1. Thus, while one model tells us that a higher share of heavy vehicles in traffic increases the number of fatalities, the other model tells us exactly the opposite. In short, the models have not been able to estimate the true effect of the independent variables on the number of fatalities. Merely by replacing one independent variable by another, all coefficients for the variables common to both models changed sign.

No substantive interpretation of the models is possible. The regression coefficients, although precise, make no sense. If each model is considered in isolation, it may well be accepted since it to a large extent satisfies formal criteria of model quality. But if one were to apply the regression coefficients to estimate the partial effects of each variable, the results would be diametrically opposite for models 1 and 2 and impossible to interpret.

8 A discussion of alternative modelling strategies

Trying to estimate a model containing six variables when there are only 17 observations might seem hopeless. There are not enough degrees of freedom left to reliably estimate regression coefficients. Clearly, this could explain why the coefficients for the five variables that were common to the two models changed sign when the sixth variable was replaced. Nevertheless, as noted before, very many variables influence the number of traffic fatalities and there is a desire to know about the effects of as many variables as possible. The simple exercise reported above thus reflects the nature of the problem facing analysts.

What are the main strategies for developing explanatory models when there are very many potentially explanatory variables? There are two main options. The first is to increase the number of observations by extending the analysis from a single country to many countries. Page (2001) created data set for 21 countries for 1980-1994, generating a total of 21 x 15 = 315 observations. Such a data set is referred to as a panel data set and contains both variation between countries (cross-sectional variation) and over time. He estimated a model including seven independent variables, noting that many potentially important variables were not included. It is clear that the model had very large residual terms; for some countries the model-estimated number of traffic fatalities was less than half the actual number. Moreover, the model did not describe the decline in traffic fatalities from 1980 to 1994 very well. A majority of the residuals were negative for 1980-82, meaning that the model underestimated the number of fatalities. For 1992-94, nearly all residuals were positive, meaning that the model overestimated the number of fatalities. Thus, the model estimated a much smaller reduction in the number of traffic fatalities than actually took place. In general, limited data are available at an international level, which means that any model developed for many countries will omit many important variables and is likely to contain omitted variable bias of an unknown magnitude.

A second option is to create subgroups of traffic fatalities and identify the factors most likely to influence each group (Stipdonk and Berends 2008). A paper by Stipdonk and Berends illustrates this approach. They identified six groups and showed that the development over time of the number of fatalities differed between these groups. Using a specific group of fatalities as dependent variable ought to make more precise analyses possible. One would, for example, expect increased seat belt wearing to contribute to a decline in car occupant fatalities, but not influence pedestrian or cyclist fatalities. Thus, identifying groups of fatalities permits using the casualty subset test described by Fridstrøm (2015).

A drawback of studying groups of fatalities is that the number of fatalities in some groups may become so low that random variation makes a major contribution to annual changes. In Norway, for example, the mean annual number of fatalities involving moped or motorcycle riders from 2009 to 2014 was 23, fluctuating between 29 and 17.

9 General discussion

Why have many highly motorised countries been able to reduce the number of traffic fatalities in the past 45-50 years by around 80 percent? Are there any lessons to learn here for countries that are fast motorising and experiencing an increase in the number of fatalities? These important questions are very difficult to answer.

The principal difficulty is that very many factors influence the number of traffic fatalities. The number of variables whose effects one would like to determine exceeds the number of years during which there has been a tendency for the number of fatalities to decline. Moreover, many of these variables are almost perfectly correlated with time. In any model that includes a trend term, the effects of these variables cannot be estimated and ends up in the trend term. Omitting variables is not a good solution. It may seem to reduce problems related to co-linearity, but is likely to introduce omitted variable bias, which means that the effects estimated for the variables that are included in a model will be biased and partly reflect one or more omitted variables. The sheer number of relevant variables makes it difficult to believe that this problem can be avoided.

Can a few explanatory variables be selected from the many that are relevant? Is it possible to identify the variables that may have been most important in explaining the decline in the number of fatalities? There hardly seems to be any well-developed theoretical or empirical foundations for making such a selection. Seat belts, for example might be a candidate; it can reasonably be argued that it has saved more lives than most other road safety measures. Yet, when seat belt use was included in a parsimonious model for Norway, including just seat belt wearing, traffic tickets per million vehicle kilometres and unemployment, the seat belt variable was found to contribute to reducing both car occupant fatalities, pedestrian and cyclist fatalities and moped and motorcycle fatalities. This makes little sense, but it is easy to see why one gets this result. Over time seat belt wearing has increased, while traffic fatalities have gone down in all the three groups (car occupants, pedestrians and cyclists and moped and motorcycle riders). Hence, the variables happen to be negatively correlated. A negative regression coefficient is estimated. It makes some sense for car occupants, but hardly for the other two groups.

It is likely that very many variables happen to be correlated this way. This may generate lots of non-sensical regression coefficients in multivariate models designed to explain the decline in traffic fatalities. It does not help to assess the goodness of the models according to the usual formal criteria. The models may fit the data extremely well, and have well-behaved residual terms, yet make no sense from a substantive point of view.

10 Conclusions

The main conclusion of the study presented in this paper can be summarised as follows:

1. Very many variables influence the number of traffic fatalities. It is impossible to include all of them in a multivariate model designed to explain the decline in the number of traffic fatalities in many highly motorised countries.
2. The variables influencing the number of traffic fatalities tend to be highly correlated among themselves and with time. This makes it almost impossible to reliably estimate the effect of each variable.
3. No firm guidelines exist for selecting a limited number of variables for inclusion in an analysis. Including only a few variables is highly likely to lead to omitted variable bias.
4. Models that appear to be good according to formal criteria like goodness-of-fit and characteristics of the residual terms may contain non-sensical regression coefficients.

References

Allsop, R. E., Sze, N. N., Wong, S. C. 2011. An update on the association between setting quantified road safety targets and road fatality reduction. Accident Analysis and Prevention, 43, 1279-1283.

Christophersen, A. S., Mørland, J., Stewart, K., Gjerde, H. 2016. International trends in alcohol and drug use among motor vehicle drivers. Forensic Science Review, 28, 37-66.

Elvik, R. 2005. Has progress in improving road safety come to a stop? Report 792. Oslo, Institute of Transport Economics.

Elvik, R. 2009. The Power Model of the relationship between speed and road safety. Update and new estimates. Report 1034. Oslo, Institute of Transport Economics.

Elvik, R. 2012. Speed limits, enforcement, and public health consequences. Annual Review of Public Health, 33, 225-238.

Elvik, R. 2013. A re-parameterisation of the Power Model of the relationship between the speed of traffic and the number of accidents and accident victims. Accident Analysis and Prevention, 50, 854-860.

Elvik, R. 2015A. Chapter 3 (43-142) in: Why does road safety improve when economic times are hard? Paris, International Traffic Safety Data and Analysis Group (IRTAD – a branch of OECD).

Elvik, R. 2015B. Can electronic stability control replace studded tyres? Accident Analysis and Prevention, 85, 170-176.

Elvik, R. 2016A. Does the influence of risk factors on accident occurrence change over time? Accident Analysis and Prevention, 91, 91-102.

Elvik, R. 2016B. Association between increase in fixed penalties and road safety outcomes: A meta-analysis. Accident Analysis and Prevention, 92, 202-210.

Fridstrøm, L. 1999. Econometric models of road use, accidents, and road investment decisions. Volume II. Report 457. Oslo, Institute of Transport Economics.

Fridstrøm, L. 2015. Disaggregate accident frequency and risk modelling. A rough guide. Report 1403. Oslo, Institute of Transport Economics.

Fridstrøm, L., Ifver, J., Ingebrigtsen, S., Kulmala, R., Krogsgård Thomsen, L. 1995. Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. Accident Analysis and Prevention, 27, 1-20.

Gjerde, H., Normann, P. T., Christophersen, A. S., Samuelsen, S. O., Mørland, J. 2011. Alcohol, psychoactive drugs and fatal road traffic accidents in Norway: a case-control study. Accident Analysis and Prevention, 43, 1197-1203.

Gjerde, H., Christophersen, A. S., Normann, P. T., Mørland, J. 2013. Associations between substance use among car and van drivers in Norway and fatal injury in road traffic accidents: A case-control study. Transportation Research Part F, 17, 134-144.

Hauer, E. 2015. The art of regression modelling in road safety. New York, Springer.

Høye, A. 2011. The effects of Electronic Stability Control (ESC) on crashes - an update. Accident Analysis and Prevention, 43, 1148-1159.

Høye, A. 2014. Utvikling av ulykkesmodeller for ulykker på riks- og fylkesvegnettet i Norge. Rapport 1323. Oslo, Transportøkonomisk institutt.

Høye, A. (2015A). Safety effects of section control - An empirical Bayes evaluation. Accident Analysis and Prevention, 74, 169-178.

Høye, A. (2015B). Safety effects of fixed speed cameras - An empirical Bayes evaluation. Accident Analysis and Prevention, 82, 263-269.

Høye, A. 2016. How would increasing seat belt use affect the number of killed or seriously injured light vehicle occupants? Accident Analysis and Prevention, 88, 175-186.

Høye, A., Elvik, R., Vaa. T., Sørensen, M. W. J. 2015. The Handbook of Road Safety Measures. Web edition (in Norwegian). Oslo, Institute of Transport Economics.

Langeland, P. A., Phillips, R. O. 2016. Tunge kjøretøy og trafikkulykker. Norge sammenlignet med andre land I Europa. Rapport 1494. Oslo, Transportøkonomisk institutt.

Muskaug, R. 1980. Ulykker og andre data for 16 boligområder i Oslo. Arbeidsdokument av 15.7.1980, prosjekt 4753 risiko ved ulike reisemåter. Oslo, Transportøkonomisk institutt.

Odberg, T. A. 1996. Erfaringer med rundkjøringer i Vestfold. Hovedoppgave i samferdselsteknikk. Trondheim, NTNU, Institutt for samferdselsteknikk.

Page, Y. 2001. A statistical model to compare road mortality in OECD countries. Accident Analysis and Prevention, 33, 371-385.

Ragnøy, A. 2004. Endring av fartsgrenser. Effekt på kjørefart og ulykker. Rapport 729. Oslo, Transportøkonomisk institutt.

Sakshaug, K. 1986. Fartsgrenseundersøkelsen -85. Detaljerte resultater fra fartsdelen og ulykkesdelen. Notat 535/86 og 536/86. Trondheim, SINTEF Samferdselsteknikk.

Stipdonk, H., Berends, E. 2008. Distinguishing traffic modes in analysing road safety development. Accident Analysis and Prevention, 40, 1383-1393.

Tran, T. 1999. Vegtrafikkulykker i rundkjøringer – 1999. TTS rapport 2, 1999. Vegdirektoratet, Transport- and trafikksikkerhetsavdelingen, Oslo.

Wanvik, P. O. 2009. Road lighting and traffic safety. Do we need road lighting? Doctoral theses at NTNU 2009:66, Trondheim, NTNU.