TØI Report
539/2001

Quantified Road Safety Targets

An Assessment of Evaluation Methodology

Rune Elvik

ISSN 0802-0175
ISBN 82-480-0223-3

Oslo, October 2001

Title:	Quantified road safety targets: An assessment of evaluation methodology	**Tittel:**	Tallfestede trafikksikkerhetsmål: en metodestudie
Author(s):	Rune Elvik	**Forfattere:**	Rune Elvik
Date:	10.2001	**Dato:**	10.2001
TØI report:	539/2001	**TØI rapport:**	539/2001
Pages	74	**Sider**	74
ISBN Paper:	82-480-0223-3	**ISBN Papir:**	82-480-0223-3
ISBN Electronic:		**ISBN Elektronisk:**	
ISSN	0808-1190	**ISSN**	0808-1190
Financed by:	Research Council of Norway	**Finansieringskilde:**	Norges Forskningsråd
Project:	2581 - Using meta-analysis to summarise knowledge in transport research	**Prosjekt:**	2581 - Bruk av meta-analyse til kunnskapsoppsummering i transportforskning
Project manager:	Rune Elvik	**Prosjektleder:**	Rune Elvik
Quality manager:	Marika Kolbenstvedt	**Kvalitetsansvarlig:**	Marika Kolbenstvedt
Key words:	Before-and-after study Evaluation study Multivariate analysis Quantified target Road safety	**Emneord:**	etterundersøkelse Evalueringsstudie Før Multivariat analyse Tallfestet mål Trafikksikkerhet

Summary:

The report presents a study designed to evaluate the effects on road safety performance of quantified road safety targets set by national or local governments. Effects of such targets have been evaluated by means of both before-and-after studies and multivariate analyses. Regrettably neither of these studies could be implemented in a sufficiently rigorous manner to allow for any substantive conclusion. The main conclusion is that the study is inconclusive.

Sammendrag:

Rapporten inneholder en studie av virkninger på trafikksikkerheten av tallfestede mål for reduksjon av antallet skadde og drepte i trafikken, satt av nasjonale eller lokale myndigheter i ulike land. Virkningene av tallfestede mål er undersøkt både med før-og-etterundersøkelser og multivariat analyse. Dessverre har ingen av disse undersøkelsene latt seg gjennomføre på en så holdbar måte at det kan trekkes noen konklusjoner om virkningene av tallfestede trafikksikkerhetsmål.

Language of report: English

Institute of Transport Economics
Gaustadalleen 21, 0349 Oslo, Norway
Telefon 22 57 38 00 - www.toi.no

Transportøkonomisk Institutt
Gaustadalleen 21, 0349 Oslo
Telefon 22 57 38 00 - www.toi.no

Preface

This report presents an evaluation of the effects on road safety performance of quantified road safety targets set by national or local governments. The study was prompted by the invitation of the author, Rune Elvik, to give a lecture on the effectiveness of quantified road safety targets at the first “Best in Europe” road safety conference in Brussels in September 2000. The author and the Institute of Transport Economics wishes to thank the following persons who have contributed to this study:

Executive Director Jeanne Breen of the European Transport Safety Council for taking the initiative to conducting this study.

Senior research officers Axel Elsner and Suzanne Reichwein of the Bundesanstalt für Strassenwesen in Germany for generously providing data from the IRTAD database.

Senior research officer Peter Christensen of the Institute of Transport Economics, for performing the multivariate analyses presented in this report, and otherwise serving as an invaluable discussion partner.

Neither of these individuals are responsible for the contents of the report, nor any errors it might contain.

Work on this report has in part been funded by the Research Council of Norway by means of a strategic research programme on meta-analysis, in part by the Institute of Transport Economics by means of support granted to Rune Elvik in his capacity as Chief Research Officer.

Oslo, October 2001
INSTITUTE OF TRANSPORT ECONOMICS

Knut Østmoe — *Marika Kolbenstvedt*
Managing director — Head of department

Table of contents

Sammendrag

Summary

1 Introduction..1

2 Previous research – hypotheses about the effects of quantified road safety targets...2

2.1 Previous research 2

2.2 Hypotheses...4

3 Quantified road safety targets in the OECD-countries..................................6

3.1 National targets 6

3.2 Local targets...7

4 Design of evaluation study 9

4.1 Aims and challenges of road safety evaluation studies 9

4.2 Sources of data – variables included...10

4.3 Study design considerations for before-and-after studies............................10

4.4 Study design considerations for multivariate analyses 13

4.4.1 Choice of variables to be included...13

4.4.2 Functional form of relationships between variables.............................14

4.4.3 Assumptions regarding residual terms...14

4.4.4 Choices made for the design of multivariate analyses..........................14

4.5 Statistical analysis and synthesis of results of before-and-after studies......15

4.6 Statistical analysis and synthesis of results of multivariate analyses 18

4.7 Deductions from main hypothesis 18

5 Results...21

5.1 Before-and-after study of national road safety targets 21

5.1.1 Australia...21

5.1.2 Denmark...22

5.1.3 Finland 23

5.1.4 France 29

5.1.5 Hungary 30

5.1.6 Iceland...31

5.1.7 The Netherlands...32

5.1.8 New Zealand...34

5.1.9 Norway...35

5.1.10 Poland 38

5.1.11 Spain 39

5.1.12 Sweden..40
5.1.13 United Kingdom..42
5.1.14 The United States of America..43
5.2 Before-and-after study of local road safety targets..44
5.2.1 Local road safety targets in Norway..44
5.2.2 Local road safety targets in Australia..45
5.2.3 Local road safety targets in Germany..45
5.3 Synthesis of results of before-and-after studies..46
5.3.1 Testing the consistency and homogeneity of the results of the before-and-after studies..46
5.3.2 Meta-analysis of effects of quantified road safety targets..47
5.4 Multivariate analyses of effects of quantified road safety targets..51
5.4.1 Overview of models fitted..51
5.4.2 Results by model..53
5.4.3 Synthesis of results of multivariate analyses..55
5.5 Synthesis of results of before-and-after studies and multivariate analyses..55
6 Discussion..57
6.1 Two ways of interpreting study findings..57
6.2 History repeats itself until it suddenly no longer does so..57
6.3 Alternative analyses of before-and-after studies..58
6.4 Methodological issues in multivariate analyses..61
6.5 Internal validity of study findings (causal interpretation)..63
6.5.1 Existence of statistical relationship..64
6.5.2 Strength of statistical relationship..64
6.5.3 Internal consistency of statistical relationship..65
6.5.4 Clarity with respect to direction of causality..65
6.5.5 Control of confounding factors..67
6.5.6 Presence of a dose-response pattern..67
6.5.7 Specificity of effect in subset of data..68
6.5.8 Knowledge of causal mechanisms..68
6.5.9 Results reproduced in similar studies..69
6.5.10 Conclusions with respect to internal validity..69
7 Conclusions..71
8 References..72

TØI rapport 539/2001
Forfatter: Rune Elvik
Oslo 2001, 74 sider

Sammendrag:

Tallfestede trafikksikkerhetsmål: en metodestudie

Denne rapporten inneholder en studie av virkninger av tallfestede trafikksikkerhetsmål. Med tallfestede trafikksikkerhetsmål menes mål for reduksjon av antallet skadde og drepte i trafikken som er satt av myndighetene i ulike land. Eksempelvis har Sverige et mål på høyst 270 drepte i vegtrafikkulykker i år 2007.

Tallfestede trafikksikkerhetsmål er de siste årene satt i en rekke land. De fleste land i Vest Europa, samt USA, Australia og New Zealand har nå satt tallfestede mål for bedring av trafikksikkerheten og utarbeidet handlingsprogrammer med sikte på å nå disse målene. Det er derfor også økende interesse for hvilke erfaringer man har høstet med slike mål. Hovedproblemstillingen som undersøkes i denne rapporten er:

Bidrar det å sette tallfestede trafikksikkerhetsmål til at et land, eventuelt en del av et land, oppnår bedre trafikksikkerhet enn man ellers ville ha gjort?

For å besvare dette spørsmålet, er erfaringer med 22 tallfestede mål satt av nasjonale myndigheter i 15 land, og 13 lokale tallfestede mål, satt av lokale myndigheter i 3 land, undersøkt. Undersøkelsen er utført med to metoder: En før-og-etterundersøkelse, og en multivariat analyse. Den multivariate analysen omfatter bare mål satt av nasjonale myndigheter.

Det ble formulert seks hypoteser om virkninger av tallfestede mål. Fire av disse hypotesene støttes av før-og-etterundersøkelsen, den femte forkastes. Resultatene er tvetydige for den sjette hypotesens vedkommende. Før-og-etter undersøkelsen tyder på at land eller lokale myndigheter som har satt tallfestede mål oppnår litt bedre trafikksikkerhet enn land eller lokale myndigheter som ikke har satt slike mål. Forskjellen er imidlertid svært liten, mindre enn 1% reduksjon av antallet drepte i trafikken per år i den perioden et tallfestet mål har vært virksomt.

De multivariate analysene gir ikke grunnlag for entydige konklusjoner. Det er overveiende sannsynlig at resultatene av de multivariate analysene er sterkt påvirket av utelatte variabler og kollinearitet mellom de variabler som er inkludert i analysene.

Undersøkelsen ble lagt opp med sikte på å utelukke flest mulig av de mange feilkilder og metodiske vansker man står overfor i studier der utviklingen av trafikksikkerheten over tid sammenlignes mellom ulike land. Det har dessverre likevel ikke lykkes å gjennomføre undersøkelsen på en slik måte at man kan utelukke at ulike feilkilder i datagrunnlaget og/eller metoden forklarer resultatene. Hovedkonklusjonen er derfor at undersøkelsen ikke gir et holdbart grunnlag for å si noe om virkninger av tallfestede

trafikksikkerhetsmål. Så vel før-og-etterundersøkelsene som de multivariate analysene støtte på en rekke problemer som ikke lot seg løse tilfredsstillende.

Det mest alvorlige problemet man møter i denne typen forskning er at det er relativt få studieobjekter – kanskje 15-20 land – og svært mange faktorer, mange av dem ukjente, som påvirker trafikksikkerheten for hvert studieobjekt. Man kan bare få kontrollert for virkningen av noen få av disse mange faktorene. I før- og etterundersøkelsen er hvert land med mål sammenliknet med et land uten.

Selv om undersøkelsen i utgangspunktet var ment å gi et svar på spørsmålet om hvordan tallfestede trafikksikkerhetsmål virker, utviklet den seg til å bli en metodestudie. Den viktigste lærdommen er en økt erkjennelse av hvor vanskelig det er å si noe holdbart om virkninger av en bestemt faktor – i dette tilfellet hvilket mål myndighetene setter for trafikksikkerheten – i det mylder av faktorer som påvirker trafikksikkerheten.

Hovedpunktene i metodedrøftingen som er gjennomført i rapporten, er oppsummert i tabell S.1.

Tabell S.1: Oppsummering av drøfting av resultatenes interne validitet

Kriterium	Vurdering
1: Statistisk sammenheng mellom variabler	Statistisk signifikant nedgang i antall drepte påvist i før-og-etter undersøkelsen, men ikke i den multivariate analysen
2: Styrken I statistisk sammenheng	Svak sammenheng både I før-og-etterundersøkelsen og den multivariate analysen
3: Intern konsistens i statistisk sammenheng	67% av enkeltresultatene i før- og etterunder-søkelsen peker i same retning som hovedresultatet
4: Entydighet i årsaksretning	Det er ikke mulig å utelukke en omvendt årsaks-sammenheng – at land som gjør det bra velger å sette mål, snarere enn at de gjør det bra fordi de har satt mål
5: Kontroll for bakenforliggende variabler	Mangelfull både i før-og-etterundersøkelsen og i de multivariate analysene
6: Forekomst av dose-responsmønster i resultatene	Kan påvises i før-og-etterundersøkelsen, men ikke i de multivariate analysene
7: Virkning påvises bare i tiltakets målgruppe	Dette kriteriet har ikke latt seg anvende
8: Beskrivelse av virkningsmekanisme	En slik beskrivelse har ikke latt seg gjennomføre – det er uvisst hvilke handlingsprogrammer ulike mål har ført til
9: Tilsvarende funn i andre undersøkelser	Det finnes få andre undersøkelser, og de få som finnes er metodisk ganske svake

Man kan av denne undersøkelsen ikke trekke den slutning at det ikke har noen hensikt å sette et tallfestet mål for bedring av trafikksikkerheten. Mange av de mål som er satt er blitt nådd og har i den forstand virket etter hensikten.

TØI report 539/2001
Author: Rune Elvik
Oslo 2001, 74 pages English language

Summary:

Quantified Road Safety Targets: An Assessment of Evaluation Methodology

This report contains an evaluation of the effectiveness of quantified road safety targets. A quantified road safety target is any numerical target set by national or local governments for the reduction of the number of people who are killed or injured in road traffic accidents. As an example, Sweden has set a target of not more than 270 road accident fatalities in 2007.

Quantified road safety targets have been set in a number of countries in recent years. Most countries in Western Europe, as well as the United States of America, Australia, and New Zealand have set quantified targets for improving road safety and developed programmes designed to realise these targets. There is, accordingly, an increasing international interest in learning about the effectiveness of quantified road safety targets in improving road safety performance. The main question, which is investigated in this report is:

Do quantified road safety targets contribute to improving the road safety performance of countries, or local governments, that have set such targets?

In order to answer this question, the effects of 22 targets set by national governments in 15 countries, and 13 targets set by local governments in 3 countries have been evaluated. The evaluation study employed two research designs: A before-and-after study, and a multivariate analysis. The latter analysis comprised targets set by national governments only.

Based on previous research, six hypotheses about the effects of quantified road safety targets were proposed.

The before-and-after studies relied on selecting a comparison country for each country that had set a target. The results of the before-and-after supported four of the six hypotheses. One hypothesis was rejected. Evidence was inconclusive with respect to the sixth hypothesis. On the average, having a quantified road safety target was associated with a very small gain in safety performance, less than 1% per year. The multivariate analyses were inconclusive. It is highly likely that these analyses have been influenced by omitted variable bias, as well as instability due to collinearity among the explanatory variables.

Unfortunately, one cannot rule out the possibility that errors in data or analyses explain study findings. Although an effort was made to carry out both the before-and-after studies and the multivariate analyses rigorously, a number of difficulties, for which no satisfactory solutions were found, were encountered in both the before-and-after studies and in the multivariate analyses.

The report can be ordered from:
Institute of Transport Economics, PO Box 6110 Etterstad, N-0602 Oslo, Norway
Telephone: +47 22 57 38 00   Telefax: +47 22 57 02 90

The main conclusion of this study is, therefore, that it is inconclusive: It does not provide a credible basis for estimating the effects on safety performance of quantified road safety targets.

1 Introduction

The number of people who are killed in road accidents remains large in many countries. It is true that the number of road accident fatalities in some countries, notably the United States, Japan and many countries in Western Europe has been reduced from a peak level, which was reached around 1970. But no country has reached a level of road safety that it regards as satisfactory. In recent years, an increasing number of countries have set ambitious long term quantified targets for improving road safety. Denmark, for example, recently published a road safety programme that aims to reduce the number of road accident fatalities from 499 in 1998 to 300 in 2012 (Færdselssikkerhedskommissionen 2000). Other countries that have set targets for the reduction of road accident fatalities include Australia, Finland, France, Great Britain, the Netherlands, Sweden, and the United States.

This report examines the effectiveness of quantified road safety targets in improving road safety performance. The following questions are at the focus of the study:

- Are countries or local governments that have set quantified road safety targets more successful in bringing down the number of people killed in road accidents than countries or local governments that have not set such targets?
- Is every quantified road safety target associated with a more successful safety performance, or are some types of quantified targets more successful than others? What are the characteristics of the most successful quantified road safety targets?

Evaluating the effects of quantified road safety targets is difficult. In addition to presenting evidence that sheds light on the two main questions, the report contains an extensive discussion of methodological problems encountered in a study designed to evaluate the effects of such a target. Two different study designs have been used in the report: A before-and-after study, and a multivariate regression analysis.

2 Previous research – hypotheses about the effects of quantified road safety targets

2.1 Previous research

There has not been much previous research concerning the effects on safety performance of setting quantified targets for improving road safety.

Risser and Michalik (1987) studied the effects of a programme designed to reduce the annual number of accidents by 10% in Austria. A small, short term effect was found. After only about half a year, the programme appeared to have lost its momentum and no remaining effect could be found.

Lebrun (1989) summarises the experience gained in the "Action –10%" road safety programme implemented in France in the 1980s. This was a programme rewarding local governments who succeeded in reducing the number of injury accidents by 10% from one year to the next. During the first year of the programme, nearly all *departements* in France participated in the programme, and the majority did achieve the targeted 10% reduction of the number of accidents. In the second, third and fourth year of the programme, participation in it gradually fell, and fewer local governments were able to reduce the number of accidents by 10%. In a subsequent analysis, Jaeger and Lassarre (2000) estimated the effects of "Action –10%" by means of a multivariate Poisson-regression model belonging to the international DRAG family of road accident models. They estimated the effect of "Action –10%" to a 2% increase in injury accidents and a 3% increase in fatal accidents. Neither of these effects were statistically significant.

Schlabbach (1990) presents a before-and-after study of a one year road safety programme designed to achieve a 10% reduction of the number of accidents in the city of Darmstadt in Germany. The number of injury accidents in Darmstadt was reduced by 10.6% in the year this programme was effective. In the comparison city of Kassel, the number of accidents went down by 2.4% the same year.

Elvik (1993) compared the safety performance of Norwegian counties with and without quantified road safety targets during the 1980s. He found that ambitious quantified targets, meaning targets aiming for a drastic reduction in the number of road accidents, were associated with a better safety performance than less ambitious quantified targets, or no quantified targets at all. An OECD report (OECD Scientific Expert Group 1994) also concluded that setting quantified road safety targets appears to be associated with a better safety performance. This conclusion was based chiefly on the study of Elvik (1993).

A recent OECD report on road safety management and implementation strategies (OECD Scientific Expert group 2000) concludes that: “The setting of targets has proven its value in many countries. Target setting leads to more realistic and effective programmes, results in more integration of institutional efforts and often produces more focused allocation of resources by securing political commitment.”

The evidence presented in the report to support this conclusion is, however, not very extensive. As far as effects on the number of accidents or fatalities are concerned, the only evidence quoted is the results of the Norwegian study, referred to above. But this study says nothing about the effectiveness of targets at an international level.

In a preliminary version of the analyses that are extended in this report, Elvik (2000A; 2000B) concluded that countries with quantified road safety targets have a slightly better safety performance that countries without quantified road safety targets. The difference amounted to less than 1% per year, but was statistically significant. A total of 16 targets set by national governments and 12 targets set by local governments were included in the study. Effects of three potentially confounding variables were discussed. It was concluded that these confounding variables may have influenced changes in road safety in some of the countries that were studied. The effects of the confounding variables were, however, not estimated statistically.

The idea of setting quantified road safety targets is widely supported, both by the OECD (2000), the European Transport Safety Council (Breen 2000), and by all governments that have set such targets. It is widely believed that setting a quantified target for improving road safety, especially a demanding target, leads to a more effective road safety policy. There are, however, few evaluation studies which show that quantified road safety targets are effective in improving road safety. The studies quoted above are nearly all before-and-after studies that do not adequately control for confounding variables. This study design is particularly difficult to implement in a sufficient rigorous manner when the objective is to determine how targets contribute to the safety performance of different countries.

It is notoriously difficult to compare changes in road safety over time between countries, let alone determine what explains these changes. Nevertheless, trying to make the best use of whatever evidence available data provides, is better than concluding at the outset that research is impossible. There is a distinct possibility that, in the end, one has to conclude that research attempting to evaluate the effectiveness of quantified road safety targets at the international level is inconclusive. On the other hand, one should not forget that the alternative to imperfect research in this case is not “perfect” research. It is impressions or opinions based on flimsy evidence, which is unlikely to have been analysed very well, if at all. It is in everyone’s interest to try to assess the effectiveness of quantified road safety targets set by national governments as rigorously as possible, even if a study which is ideal from a methodological point of view cannot be performed.

2.2 Hypotheses

Based on previous research and common sense, six hypotheses are proposed concerning the effects of quantified road safety targets:

Hypothesis 1:

Countries or local authorities that have set quantified road safety targets have a better road safety performance than countries or local authorities that have not set quantified road safety targets.

Setting a safety target signals a concern about road safety and an intention to improve it. Once a target has been set, effective ways of realising it must be found. This generates an interest in setting efficient priorities for road safety measures. One would therefore expect countries or local governments that set quantified road safety target to adopt more effective road safety programmes than they would otherwise have done.

Hypothesis 2:

Highly ambitious quantified road safety targets are associated with a better road safety performance than less ambitious road safety targets.

This hypothesis is based on the findings of the study reported by Elvik (1993). An ambitious target is one that aims for a drastic reduction of the number of road accident fatalities and injuries.

Hypothesis 3:

Long-term quantified road safety targets are associated with a better road safety performance than short-term quantified road safety targets.

A long-term target gives more time to implement road safety measures than a short-term target. Moreover, the basis for setting and assessing a short-term target may be unduly influenced by random fluctuations in the number of accidents (an exceptionally bad year is likely to be followed by a more normal year).

Hypothesis 4:

Quantified road safety targets set at the national level of government are associated with a better safety performance than quantified road safety targets set at the regional or local levels of government.

National governments have at their disposal more safety measures than local governments, especially with respect to legislation and vehicle safety standards.

Hence, national governments can take stronger action to improve road safety than local governments.

Hypothesis 5:

The first generation of quantified road safety targets is associated with a better road safety performance than successive generations of quantified road safety targets.

Previous research (Elvik 1993) indicates that quantified road safety targets tend to be revised in the direction of caution, especially when the original target was not realised. In addition, the first time a quantified road safety target is set, it may act as a more powerful motivating force than on successive occasions.

Hypothesis 6:

Targets set by governments in countries with a bad safety record (a high number of road accident fatalities compared to other countries at approximately the same level of motorization) are more successful than targets set by governments in countries with a comparatively good safety record.

In countries that have a high incidence of road accidents, the scope for improvement is likely to be greater than in countries that have already succeeded in bringing down the number of accidents. Although all countries can improve road safety, the largest improvements can be made in countries that have a high fatality rate (per inhabitant, adjusted for motorization rate).

These hypotheses imply a certain pattern in the results of a study designed to evaluate the effectiveness of quantified road safety targets. If the hypotheses are supported, this will be taken as an indication that study findings primarily reflect the effects of road safety targets, not of confounding factors that were not controlled in the study.

The hypotheses do not, however, constitute a very strong theoretical basis for the study – being based as they are mostly on the rather scant previous research in this area and on common sense.

3 Quantified road safety targets in the OECD-countries

3.1 National targets

Information about quantified road safety targets at the national level of government in the OECD-countries has been taken from a number of sources.

Targets set by national governments were identified on the basis of previous research (Elvik 1993; OECD Scientific Expert Group 1994) and current policy documents (Færdselssikkerhedskommissionen 2000; Hungarian government 1993; Ministry of Transport and Communications Finland 1997; Land Transport Safety Authority 1995; National Road Safety Committee 2000; OECD Scientific Expert Group 2000; US Department of Transportation 2000; Transport Canada 2001; Technical University of Gdansk 1996). Table 1 lists all national road safety targets that have been identified. The targets are listed by country. Countries are listed alphabetically, successive targets set in each country are listed chronologically.

Table 1: Quantified road safety targets set by national governments in the OECD-countries.

Country	Target number (in country)	Base year or years	Target year	Number of fatalities in base year	Targeted number of fatalities	Percent annual change
Australia	1	1992	2001	1974	1930	-0.3
	2	1997	2005	1767	1600	-1.2
Denmark	1	1986-88	2000	711	427	-4.2
	2	1998	2012	499	300	-3.6
Finland	1	1972	1979	1156	578	-9.4
	2	1976	1980	804	550	-9.2
	3	1978	1989	610	325	-5.6
	4	1980	1989	550	300	-6.6
	5	1986	1994	612	370	-6.2
	6	1988	2000	734	370	-6.1
France	1	1997	2002	8444	4222	-12.9
Hungary	1	1992	2000	2101	1575	-3.5
Iceland	1	1991-96	2000	19	15	-5.4
Netherlands	1	1985	2000	1438	1079	-1.9
	2	1986	2010	1527	764	-2.8

Table 1: Quantified road safety targets set by national governments in the OECD-countries, continued.

Country	Target number (in country)	Base year or years	Target year	Number of fatalities in base year	Targeted number of fatalities	Percent annual change
New Zealand	1	1990	1994	729	650	-2.9
	2	1990	2001	729	420	-4.4
	3	1999	2010	509	295	-4.5
Norway	1	1984-86	1993	420	420	0
Poland	1	1991	2001	7901	6000	-2.3
Spain	1	1992	1999	7818	5473	-5.0
Sweden	1	1989	2000	904	400	-7.2
	2	1996	2007	537	270	-6.1
United Kingdom	1	1981-85	2000	5793	3862	-2.6
	2	1994-98	2010	3727	2236	-4.2
United States	1	1996	2008	42065	33500	-1.8

A total of 26 quantified road safety targets set in 14 countries are listed in Table 1. The table includes both targets that were set many years ago and have expired, and targets that have been set for years in the future, such as 2008 or 2010. The first quantified road safety target ever set by a national government, was the target set in 1973 in Finland to reduce the number of road accident fatalities by 50%, from 1156 in 1972 to 578 in 1979.

To evaluate the effects of targets, the minimum length of the before- and after-periods has been set to three years. The latest year for which a final count of road accident fatalities is available, is 1999. This means that the most recent targets set in Australia (for 2005), Denmark (for 2012), New Zealand (for 2010), and the United Kingdom (for 2010) cannot yet be evaluated. Both Australia, Finland, the Netherlands and Sweden set their second (or even third) target before the first target was to be realised. For these four countries, therefore, two targets have been in force at the same time.

3.2 Local targets

Few data are available concerning quantified road safety targets set by local or regional governments. The only data that have been found refer to targets set by the city of Darmstadt in Germany (Schlabbach 1990), Norwegian counties (Elvik 1993) and targets set by Australian states (OECD Scientific Expert Group 1994).

Targets set by Norwegian counties during the 1980s have been grouped according to their levels of ambition into eight homogeneous groups. Table 2 shows this grouping. The grouping is based on the level of ambition embodied in the targets, as well as the road planning term to which they applied (either the 1982-85 or 1969-89 planning terms). There are nineteen counties in Norway. Eleven counties had quantified road safety targets during the 1982-85 planning term, eight did not.

Sixteen counties had quantified road safety target during the 1986-89 planning term, three counties did not have such targets.

Table 2: Groups of quantified road safety targets set by Norwegian counties in the 1980s. Derived from Elvik 1993

Group	Base year(s)	Target year(s)	Base number of accidents	Target number of accidents
1 (-50%; 10 years)	1978-80	1987-88	1868	934
2 (-26%; 6 years)	1978-80	1985	1928	1431
3 (-18%; 5 years)	1978-80	1985	1926	1572
4 (-31%; 8 years)	1978-82	1988-89	1713	1178
5 (-12%; 7 years)	1982-83	1989-90	2010	1773
6 (-0%; 4 years)	1985	1989	1068	1068
7 (-27%; 7 years)	1982-83	1989-90	1835	1331
8 (-0%; 8 years)	1982-85	1989-90	1006	1006

The targets set by Australian states are listed in the OECD-report referred to above. The following states were included: New South Wales, Victoria, Western Australia, and Northern Territories. The targets set by these states applied to the period 1991-2000. The target set in Darmstadt was to reduce accidents by 10% during one year. An evaluation of this target has been presented by Schlabbach (1990).

4 Design of evaluation study

4.1 Aims and challenges of road safety evaluation studies

As noted in the introduction, the evaluation of quantified road safety targets in this report relies on two different study designs: a before-and-after study, and a multivariate analysis. Both these study designs have the same objective, which is to estimate, as precisely as possible, the effect of quantified road safety targets on safety performance.

It is useful to state a formal definition of the concept of "effect" of a road safety target or road safety programme: *An action intended to improve road safety has an effect if it leads to a lower expected number of accidents or a lower expected number of road accident victims than would otherwise have occurred, that is in the absence of the programme with everything else kept constant*.

In order to determine the effect of a road safety programme, one has to (Hauer 1997):

- Predict what the level of safety would have been if the programme had not been introduced,
- Estimate the effect of the programme, by comparing the actual number of accidents or accident victims to the numbers predicted to occur in the absence of the programme, and
- Rule out, as far as possible, other explanations of the changes in safety than the programme whose effects the study aims to determine (confounding factors).

These are the challenges any road safety evaluation study face. The best way to eliminate the effects of confounding factors – that is any factor that influences safety other than the safety programme of interest – is to use an experimental study design. In the present study, an experimental study design could not be applied. In a non-experimental study, the level of control of confounding factors attained is one of the most crucial aspects of study design or data analysis. Poor control of confounding factors can produce very misleading estimates of the effects of road safety programmes or road safety targets. Striking illustrations of this can be found in studies reported by Elvik (1997; 2002). The most important consideration in designing non-experimental road safety evaluation studies is, therefore, to control for as many potentially confounding factors as possible.

4.2 Sources of data – variables included

The main source of data used is the IRTAD data base, administered by the Bundesanstalt für Stassenwesen on behalf of the OECD. In addition to IRTAD, the yearly "Economic outlook" publication of OECD has been used. The following data have been extracted for each country from these sources for the years 1970-1998, or for shorter periods for countries that have not been members of IRTAD during the whole period:

1. Annual number of people killed in road accidents, converted to the 30-day definition for countries that use a different definition of a road accident fatality
2. Annual mean population in thousands
3. Registered number of motor vehicles in thousands, excluding mopeds
4. Annual percentage real growth of the gross national product. Reduction is listed as a negative growth rate

In addition to these variables, the following variables have been coded for use in multivariate analysis:

1. Dummy variables to identify countries (one dummy for each country)
2. Year (1970, 1971, … 1998)
3. Dummy variable to indicate the presence of a quantified road safety target (1 = yes, 0 = no)
4. Variable indicating the annual mean percentage reduction of the number of road accident fatalities envisaged by a quantified road safety target
5. Duration of road safety target, counting 1 for the first year it was effective, 2 for the second year, and so on.

No data was obtained on the number of people injured in road accidents in the IRTAD member countries. These data are not comparable across countries, due to different definitions of a reportable accident and varying levels of reporting in official accident statistics (Elvik and Mysen 1999).

Furthermore, no data was obtained regarding traffic volume, in million vehicle kilometres of travel. These data are available for recent years for some countries, but there are too many gaps in the data set for the data to be useful. The registered number of motor vehicles will therefore be used to indicate traffic exposure.

4.3 Study design considerations for before-and-after studies

The before-and-after study was designed according to the guidelines proposed by Ezra Hauer (1997). According to Hauer, potentially confounding factors in road safety evaluation studies can be sorted into two groups:

1. Factors which are known, can be measured, and whose effects can be estimated statistically, for example by means of a multivariate analysis
2. Factors which are unknown, cannot be named or otherwise identified, and whose effects one tries to account for by means of a comparison group.

Whether or not this notion of the role of a comparison group in before-and-after studies of road safety measures is correct or fruitful will not be discussed at this stage. An important point in non-experimental studies, is that the comparison group is *chosen*; it does not result from randomisation. Having defined the role of a comparison group, Hauer (1997) goes on to develop guidelines for the choice of comparison group among several candidates. A good comparison group has the following characteristics:

1. It is known from past accident history to accurately predict changes in accident counts in the treated group.
2. It is sufficiently large for random variation in the count of accidents not to influence year to year changes very much.
3. It is not subject to the treatment or safety programme whose effects the study tries to estimate.

The first of these characteristics is intended to serve as a criterion of similarity in terms of the effects of all causal factors on the count of accidents: If two groups behaved similarly in the past, one may take this as an indication that they have been subject to similar processes resulting in road accidents. The second characteristic is intended to minimise the contribution of a comparison group to the statistical uncertainty of any estimate of the effect of a safety programme. These two characteristics mimic the effects of randomisation in large samples, which is to make the treatment and control groups identical with respect to the effects of confounding factors. Observe that the term "control group" refers to a randomised controlled trial, whereas the term "comparison group" refers to a non-experimental study.

To illustrate the logic and workings of the first characteristic listed above, similarity of accident history, consider the case of Belgium as a comparison group for the Netherlands, which has set a quantified road safety target. Figure 1 shows the predictive performance of Belgium as comparison group for the Netherlands during the period 1973-1985. During this period, the before-period, neither country had a quantified road safety target.

The first year for which Belgium could be used to predict the number of fatalities expected to occur in the Netherlands was 1974. From 1973 to 1974, the number of fatalities in Belgium changed from 2915 in 1973 to 2665 in 1974. The number of fatalities in the Netherlands in 1973 was 3092. Using Belgium to predict the number for 1974 in the Netherlands, we get 3092 x (2665/2915) = 2827. The actual number for 1974 in the Netherlands was 2546. In figure 1, the dots shows the actual number of road accident fatalities each year in the Netherlands, the line shows the predicted number, using Belgium as a comparison group. During most years, the predicted and actual numbers are quite close, which means that Belgium performs well in predicting the Dutch experience.

The accuracy of reach prediction can be evaluated as an odds ratio, defined as follows:

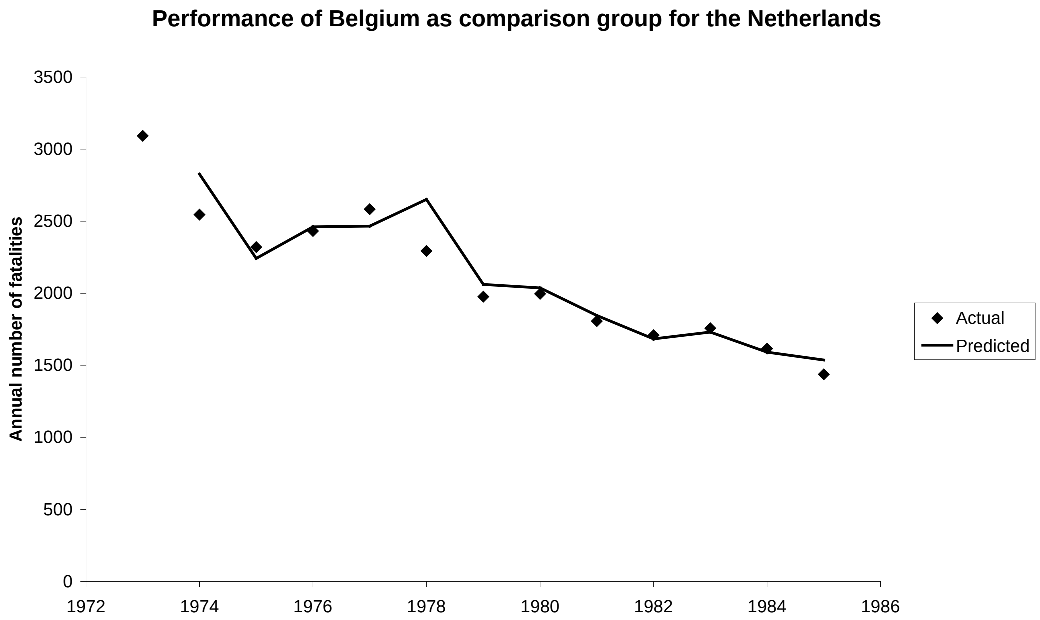

Figure 1: Predictive performance of Belgium as comparison group for the Netherlands 1973-1985

$$\text{Odds ratio} = \frac{\left(\dfrac{\text{Treated country year 2}}{\text{Treated country year 1}}\right)}{\left(\dfrac{\text{Comparison country year 2}}{\text{Comparison country year 1}}\right)}$$

If 1973 = 1 and 1974 = 2, the Netherlands is the treated country and Belgium the comparison country, the odds ratio comes to (2546/3092)/(2665/2915) = 0.901. If prediction is perfect, the odds ratio ought to be 1.000. Predictive performance can therefore be assessed by examining how close to the ideal value of 1.000 the mean odds ratio based on a set of predictions is. For the case of Belgium and the Netherlands, shown in Figure 1, twelve predictions were made. The (unweighted) mean odds ratio was 0.978. The variance of the odds ratio was 0.003. A 95% confidence interval for the mean odds ratio (standard error of the mean) goes from 0.947 to 1.009, which shows that the mean odds ratio did not differ significantly from the ideal value of 1.000.

In the before-and-after study, a comparison country has been chosen for each country with a quantified road safety target, based on the logic explained above. Target and comparison countries have been matched according to prior accident history, and not with respect to any other variables. In general, the length of the before-period was at least four years, often eight to ten years. It was not possible, within the resources available for this study, to explore the implications of varying the length of the before-period. The after period was in no case shorter than three years.

4.4 Study design considerations for multivariate analyses

The basic elements in the design of a multivariate analysis of accident data concern:

1. The set of variables to be included
2. Specification of the functional form of the relationship between each variable and the count of road accident fatalities
3. Specification of the assumptions made with respect to the residual terms of a multivariate model.

These points will be discussed in turn.

4.4.1 Choice of variables to be included

As far as choice of variables to be included is concerned, one can err by including both too few variables, too many, or by incorrect specification of the model by including variables that are endogenous in relation to the main independent variable of interest.

The first of these errors is known as *omitted variable bias*. Whether or not a statistical model has this bias, cannot be determined on the basis of purely statistical considerations. A model which does not fit the data very well, perhaps explaining very little of the variance of the dependent variable, does not necessarily suffer from omitted variable bias. The variables that ought ideally to be included in a model, have to be specified in advance, based on theoretical considerations.

The second possible error, including too many variables, can to some extent be diagnosed statistically. If some of the variables are highly correlated, this may lead to problems in estimating their effects precisely. This will show up in the form of unstable coefficients, depending on which of the highly correlated variables are included in a certain model specification. If the distribution of the residuals in a perfect model is known, one may use this as a benchmark for diagnosing a model. An *over fitted* model is one which "explains" some of the random variation of the dependent variable, in addition to the systematic part of variation. This point will be elaborated below.

A third possible error, is to *treat an endogenous variable as exogenous*, and thereby erroneously control for it. A case in point would be the following: Suppose a country sets a demanding quantified road safety target. This leads to an increase in the spending on road safety programmes. This in turn leads to fewer accidents. In a multivariate model, it would be wrong to control for the effects of spending in safety programmes, when estimating the effects of the road safety target on safety performance, since changes in spending is, itself, one of the effects of the road safety target.

4.4.2 Functional form of relationships between variables

The relationship between a pair of variables can be linear, exponential, logarithmic, hyperbolic, or any other shape that can be stated as a mathematical function. If, for example, a model assuming a linear relationship between two variables is fitted to a relationship that is in fact logarithmic, misleading estimates of effect my result. The point of modelling non-linear relationships correctly is discussed extensively by Gaudry (2000), who also gives some illustrative examples.

One can probe for the shape of the relationship between two variables by means of exploratory analysis, that is by trying different shapes, such as linear, exponential, square root, and so on. Such an approach does, however, involve data dredging, which amounts to proposing hypotheses after first looking at the data. By this approach, the real statistical reliability of the results will be unknown.

Another approach is to apply Box-Cox transformations to some or all variables, allowing the shape of their inter-relationships to be determined empirically by estimating the best fitting Box-Cox parameters. However, even this approach represents a form of data dredging, by being simply a curve-fitting device.

4.4.3 Assumptions regarding residual terms

No multivariate model can account for all variation present in the data. The analysis of accident count data is a case in which the amount and distribution the residual terms of a perfectly fitted multivariate model are known. A perfectly fitted multivariate model designed to explain accident count data leaves only random variation as unexplained. Random variation in accident counts is usually modelled by means of the Poisson distribution, an assumption which is well justified by both theoretical considerations and numerous empirical examples. In the Poisson distribution, the variance equals the mean.

If, for example, the mean number of accidents is 2, purely random variation will also be 2. Suppose the empirical variance in a data set with a mean of 2 is 5. Then, the best fit that any multivariate accident model could possibly achieve would be to explain $(5 – 2)/5 = 0.6$ or 60% of the variance. The partitioning of variance into the systematic part and the random part can be used as a benchmark to assess whether a multivariate accident model is over fitted or not (Fridstrøm et al 1995).

It is only a perfectly fitted multivariate model that can explain all systematic variation in accident counts. Most models will leave some part of the systematic variation unexplained. The distribution of residual terms is then best modelled by means of the negative binomial distribution. In large accident samples, however, both the Poisson distribution and the negative binomial distribution can be approximated by the normal distribution.

4.4.4 Choices made for the design of multivariate analyses

In order to be able to see how the design of a multivariate analysis affects its results, several analyses have been made. These analyses differ with respect to:

- The set of variables included

- The form assumed for the relationship between the variables
- The assumption made regarding the distribution of residual terms

A total of fifteen (15) analyses have been made. Details of these analyses are reported in chapter 5, which presents the results.

4.5 Statistical analysis and synthesis of results of before-and-after studies

The results of each before-and-after study of a given quantified road safety target can be laid out as a 2 x 2 table, in which the columns refers to the two periods (before or after) and the rows refer to treated or comparison country. Each 2 x 2 table produces an estimate of a quantified road safety target in terms of an odds ratio. These estimates have been analysed and combined into an overall estimate of effect by applying the log odds method for combining evidence from 2 x 2 tables (Fleiss 1981). The main features of analysis can be laid out as follows.

The changes in the number of fatalities (or, in the case of Norwegian counties, number of injured road users) associated with the adoption of a quantified road safety target were measured in terms of the odds ratio:

Effect of quantified road target = (A/B)/(C/D)

in which A is the number of fatalities in the target country in the after period, B is the number of fatalities in the target country in the before period, C is the number of fatalities in the comparison country in the after period, and D is the number of fatalities in the comparison country in the before period. Each estimate of effect was based on the total number of fatalities recorded in the whole before period and the whole after period. By pooling data for all years this way, the size of the accident sample is increased, thus making each estimate of effect more precise. On the other hand, any long term trend present in the before period will be pasted over by this approach to analysis. For the purpose of estimating the mean effect of a set of quantified road safety targets, each estimate of effect was assigned a statistical weight inversely proportional to the variance of the logarithm of the odds ratio. The variance of the logarithm of the odds ratio is:

$$v_i = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D},$$

Hence, the weight assigned to each result was:

$$w_i = \frac{1}{v_i + \eta_i}$$

in which η_i is the variance of the odds ratio describing the predictive performance of the comparison country during the before period. As noted above, the ideal mean value of this odds ratio is 1.000 with a very small variance. The weighted mean effect based on a set of g estimates is:

$$\bar{y} = \exp\left(\frac{\sum_{i=1}^{g} w_i \, y_i}{\sum_{i=1}^{g} w_i}\right)$$

Exp is the exponential function (that is 2.71828 raised to the power of the expression in parenthesis), y_i is the logarithm of each estimate of effect and w_i is the statistical weight of each estimate of effect.

There are two models for combining evidence from a set of two by two tables: the fixed effects model and the random effects model (Shadish and Haddock 1994). The weight given above refers to a fixed effects model. The fixed effects model of analysis is based on the assumption that there is only random variation in effects between the cases considered. To test the validity of this assumption, the following test statistic, Q, is estimated (Shadish and Haddock 1994):

$$Q = \sum_{i=1}^{g} w_i \, y_i^2 - \frac{\left(\sum_{i=1}^{g} w_i y_i\right)^2}{\sum_{i=1}^{g} w_i}$$

This test statistic has a Chi-square distribution with g – 1 degrees of freedom, where g is the number of estimates of effect that have been combined. If this test statistic is statistically significant, a random effects model of analysis is used. In a random effects model, the statistical weight assigned to each result is modified to include a component reflecting the systematic variation of estimated effects between cases. This component, often referred to as the variance component, is estimated as follows (Shadish and Haddock 1994):

$$\sigma_\theta^2 = \left[Q - (g-1)\right] / c$$

Q is the test statistic described above, g is the number of estimates and c is the following estimator:

$$c = \sum_{i=1}^{g} w_i - \left[\sum_{i=1}^{g} w_i^2 \Big/ \sum_{i=1}^{g} w_i\right]$$

The variance of each result now becomes:

$$v_i^* = \sigma_\theta^2 + v_i$$

The corresponding statistical weight becomes:

$$w_i^* = \frac{1}{v_i^*}$$

One should note that the variance of the logarithm of the odds ratio has been augmented twice in the random effects analysis. The first correction accounts for imperfect matching of target and comparison countries, the second correction accounts for systematic differences between countries in the effects of quantified road safety targets. A 95% confidence interval for the weighted mean estimate of effect was obtained according to the following expression:

$$\text{95\% confidence interval} = \exp\left[\left(\sum_{i=1}^{g} w_i y_i / \sum_{i=1}^{g} w_i\right) \pm 1.96 \cdot 1/\sqrt{\sum_{i=1}^{g} w_i}\right]$$

The weights in this expression are either the fixed effects weights or the random effects weights, depending on the model of analysis adopted.

To illustrate the logic of the analysis, a numerical example will be given. Finland set the first quantified road safety target in 1973. Based on its accident history during the years 1965 to 1972, Denmark was chosen as a comparison country. During the years 1966 to 1972, Denmark predicted the fatality count in Finland very well. The actual number of fatalities was 7,370. The predicted number was 7,361. The ratio of these numbers (7,370/7,361) is an index of the accuracy of prediction (1.001).

The number of fatalities in the before period in Finland was 8,419. During the after period (1973-1979), this was reduced to 5,634. The corresponding numbers in Denmark were 8,910 fatalities in the before period and 5,989 in the after period. The crude odds ratio is: (5,634/8,419)/(5,989/8,910) = 0.998, which corresponds to a net reduction of the number of road accident fatalities of 0.2% in Finland compared to Denmark. Adjusting the odds ratio by the inaccuracy of prediction (0.998/1.001), gives an adjusted odds ratio of 0.997.

The statistical weight of the crude odds ratio is: 1/(1/8,419 + 1/5,634 + 1/8,910 + 1/5,989) = 1738.442. Adjusting this weight for the variance of the odds ratio during the before period involves adding the estimate of the variance of the odds ratio to the denominator: 1/(1/8,419 + 1/5,634 + 1/8,910 + 1/5,989 + 0.008). The adjusted statistical weight becomes 116.615. Hence, a 95% confidence interval for the adjusted odds ratio is: Exp((ln((0.997)*116.615)/116.615) ± 1.96*1/√116.615).

4.6 Statistical analysis and synthesis of results of multivariate analyses

A total of fifteen multivariate analyses have been made. In principle, the results of these analyses can be synthesised the same way as the results of the before-and-after study have been. Each estimate of effect in each multivariate analysis is assigned a statistical weight inversely proportional to the variance (squared standard error) of that estimate, and results combined to form a weighted mean estimate. In practice, however, such an approach is not very meaningful, because the various multivariate models that have been tested are distinctly different in terms of the variables included, as well as the assumptions made with respect to the functional forms of relationships between variables, and the distribution of residual terms.

Rather than trying to combine formally the results of the various multivariate analyses, an attempt has been made to assess which of these analyses is the most rigorous from a methodological point of view. A most preferred model, or set of models has been identified and the results of that model will be emphasised.

Hypothesis 4 could not be tested in the multivariate analysis, because the data used for this analysis did not include targets set by local governments.

4.7 Deductions from main hypothesis

Which observations would be consistent and inconsistent with the main hypothesis of this study? This is a basic question in any empirical study. It is most easily answered by reference to the before-and-after study.

Consider the three cases shown in Figure 2. These three cases illustrate observations that are consistent and inconsistent with the main hypothesis proposed, that setting a quantified road safety target is associated with an improved road safety performance.

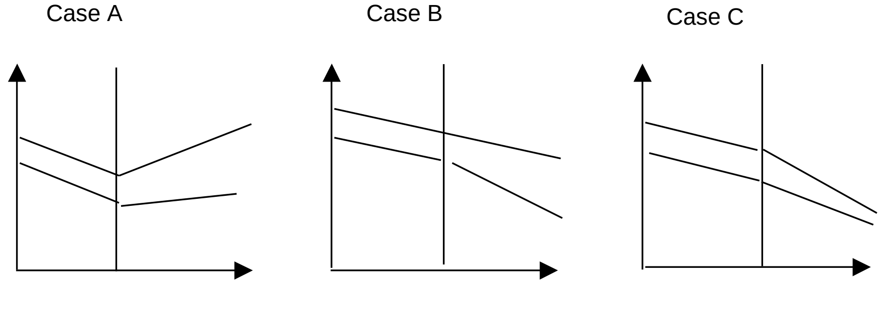

Figure 2: Three potential results of before-and-after study of quantified road safety targets

For each of the three cases, the horizontal axis shows time. A vertical line has been drawn to indicate the time at which a quantified road safety target became effective. The vertical axis shows the count of fatalities. In each cases, the uppermost curves refer to the comparison country, the lower curves refer to the target country.

Consider case A. In this case, both countries had a declining trend in the number of road accident fatalities during the before period. This reversed in the after period. However, the reversal of the trend was most pronounced for the comparison country. In case A, a simple 2 x 2 table analysis of the data would indicate that the setting of a quantified road safety target was associated with a net gain in safety performance – not because safety performance was actually better in the after period than in the before period, but simply because it deteriorated less in the target country than in the comparison country.

In case B, the target country improved its safety performance in the after period (shown by the steeper slope of curve), whereas the comparison country did not improve its safety performance (the slope of the curve remained the same as in the before period). In case B, just as in case A, a 2 x2 table analysis of the data would indicate a net gain in safety performance associated with a quantified road safety target.

In case C, the target country improved its safety performance slightly after setting a target, but the comparison country improved its safety performance even more, as indicated by the steeper slope of the upper curve in case C in the after period. Despite the fact that the target country in case C did improve its safety performance, a 2 x 2 table analysis in this case would show that the setting of a quantified road safety target was associated with a net loss in safety performance.

Of these three, hypothetical cases, it is only case B that is fully consistent with the observations one expects to make if the main hypothesis is true. There are three such observations:

1. Safety performance improves once a quantified road safety target is set. If, for example, there was a mean annual reduction of 1% in the before period and a mean annual reduction of 2% in the after period., safety performance has improved.
2. Countries with quantified targets perform better than comparison countries. This condition is fulfilled if, for example, a target country achieves a mean annual reduction of the number of road accident fatalities of 2% in the after period, whereas the comparison country only achieves a mean annual reduction of 1%.
3. The net change in safety performance from the before period to the after period should favour the target country. Suppose, for example, that the target country had an annual average fatality reduction of 1% in the before period, and 3% in the after period, a net gain of 2%. If the corresponding figures for the comparison country were 1% and 2%, respectively, the net change was 1% for the comparison country. In this case, the net change in safety performance favours the target country by 1 percentage point.

It is logically entirely possible for one or two of these points to be satisfied, while the remaining one, or the other two, fail to be satisfied. In case A in Figure 2, for example, there is a net change in safety performance favouring the target country (from, say, –2% to +1% compared to a change from –2% to +3% for the comparison country). Point 1 on the above list is, however, not satisfied.

Results that are consistent with all three points listed above lend support to the main hypothesis concerning the effect of a quantified road safety target. Results that are not consistent with all of these points will be treated as anomalous and subjected to more critical analysis.

5 Results

5.1 Before-and-after study of national road safety targets

This section presents results of the before-and-after study of national road safety targets by country. Countries are presented in alphabetic order. Multiple targets set in each country are presented in chronological order.

5.1.1 Australia

The first national quantified road safety target set in Australia was set in 1993. This target referred to the population fatality rate. A target of less than 10 fatalities per 100,000 inhabitants was set for 2001. Assuming there are 19.5 million inhabitants in Australia in 2001, this comes to less than 1930 fatalities. The base year number of fatalities (1992) was 1974. Hence the target aimed for an annual reduction of (1930/1974) = 0.9975^9 = -0.25% per year (1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001).

Figure 3 shows the annual number of traffic accident fatalities in Australia and in West Germany before and after adoption of the target in Australia.

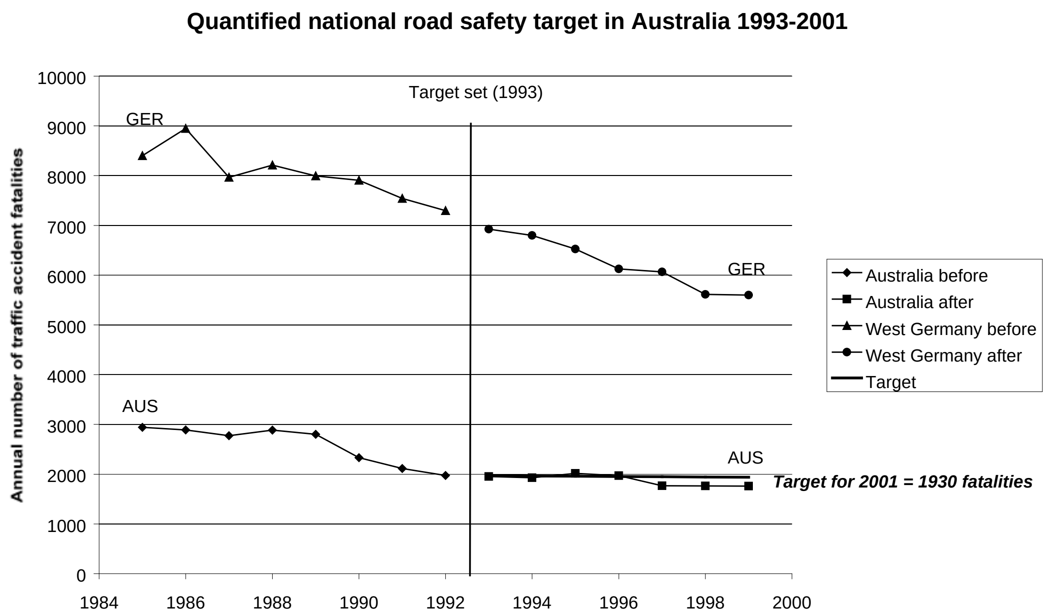

Figure 3: Australia compared to West Germany before and after adoption of a quantified road safety target in Australia

The long term trend in the number of fatalities in West Germany adequately matched that of Australia in the before period. The mean odds ratio (actual number of fatalities/predicted number of fatalities) was 0.966, with a variance of 0.006. For the years 1986 through 1992, the sum of the actual number of fatalities in Australia was 17766. The sum of the predicted number of fatalities, based on West Germany, was 18410. Table 3 provides summary data that shows the estimated effect of the quantified road safety target.

Table 3: Estimated effect of quantified road safety target in Australia 1993-2001

	Number of fatalities			Estimated effect		
Country	**Before (1985-1992)**	**After (1993-1999)**	**Target (1993-1999)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
Australia	20707	13156	13678	0.937	0.971	0.833; 1.133
West Germany	64268	43659				

The crude odds ratio is 0.937, indicated that the quantified target in Australia was associated with a net reduction in the number of fatalities of slightly more than 6% compared to West Germany. When adjusted for imperfect matching of the accident experience during the before period (Australia did perform slightly better than West Germany even then), the net effect of the target comes to slightly less than 3%. This is very far from statistical significance at the 5% level. However, the targeted number of road accident fatalities was realised.

A second quantified road safety target was set in Australia in 1997 for the year 2005. This target has been effective from 1998. It is therefore still too early to evaluate its effectiveness, as only two years of after-data are available at the moment.

5.1.2 Denmark

Just like Australia, Denmark has recently adopted a second generation road safety target. At this time, however, only the first target, adopted in 1988 for the year 2000 can be evaluated. Figure 4 and Table 4 present this evaluation.

Switzerland was chosen as comparison country. The year-to-year changes in the number of fatalities in the two countries did not match very well in all years of the before period. On the average, however, the odds ratio was 1.025, with a variance of 0.022. Analysis shows that safety performance in Denmark after adoption of the quantified road safety target was no better than in Switzerland. In fact, Switzerland did slightly better than Denmark, but the difference was not statistically significant at the 5% level.

Table 4: Estimated effect of quantified road safety target in Denmark 1989-2000

	Number of fatalities			Estimated effect		
Country	**Before (1977-1988)**	**After (1989-1999)**	**Target (1989-1999)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
Denmark	8657	6176	6262	1.176	1.165	0.869; 1.563
Switzerland	13134	7967				

Denmark did reduce the number of road accident fatalities below the targeted number, but was nevertheless outperformed by Switzerland.

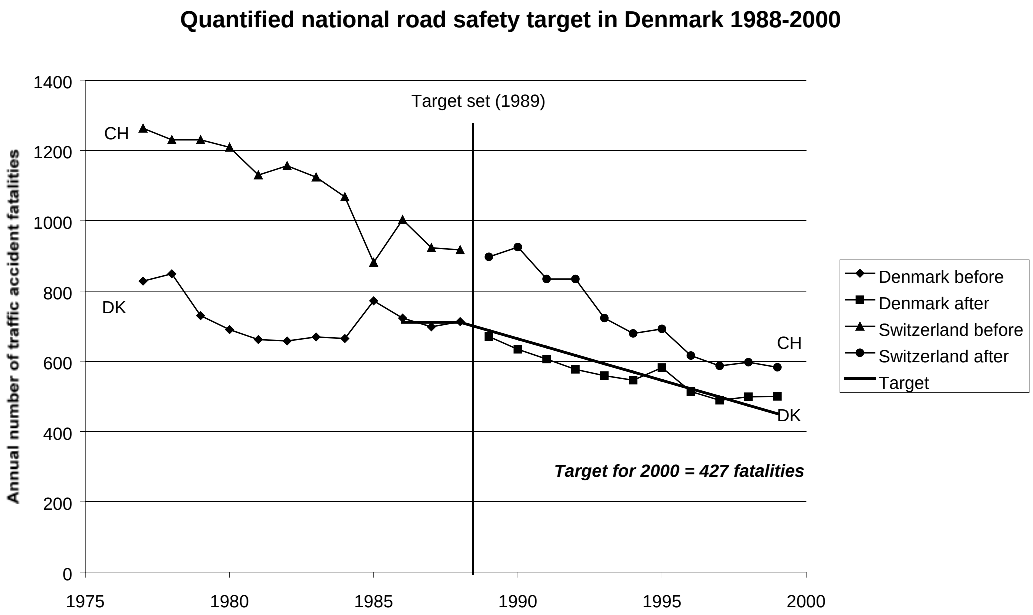

Figure 4: Denmark compared to Switzerland before and after adoption of a quantified road safety target in Denmark

5.1.3 Finland

Six successive quantified road safety targets have been evaluated for Finland. Some of these targets have overlapped in time, that is two targets have been in force at the same time, but referring to different base years or different years for target fulfilment.

Figure 5: Finland compared to Denmark before and after adoption of a quantified road safety target in Finland 1973-1979

Figure 5 and Table 5 present the evaluation of the first of these six targets. Denmark was chosen as comparison country. Denmark matched Finland very well in the before period, with a mean odds ratio of 1.003, with a variance of 0.008.

Table 5: Estimated effect of quantified road safety target in Finland 1973-1979

	Number of fatalities			Estimated effect		
Country	Before (1965-1972)	After (1973-1979)	Target (1973-1979)	Crude odds ratio	Adjusted odds ratio	95% CI
Finland	8419	5634	5782	0.998	0.997	0.831; 1.195
Denmark	8930	5989				

The adjusted odds ratio is very close to 1.000, indicated that there was no difference at all between Finland and Denmark in safety performance during the period 1973-1979, compared to the period 1965-1972. Finland realised the target that was set, but did not perform better than Denmark did in the same period.

The next target set in Finland, was a short term target for the year 1980. It has been evaluated using Denmark as comparison. Figure 6 and Table 6 contain the results of the evaluation.

Figure 6: Finland compared to Denmark before and after adoption of a quantified road safety target in Finland 1977-1980

Denmark once more matched Finland very well as a comparison country. The mean odds ratio in the before period was 1.001, with a variance of 0.027. As indicated by Table 6, Finland realised the target and performed about 20% better than Denmark in reducing the number of fatalities during the after period. The difference in safety performance was, however, not statistically significant at the 5% level.

Table 6: Estimated effect of quantified road safety target in Finland 1977-1980

	Number of fatalities			Estimated effect		
Country	**Before (1973-1976)**	**After (1977-1980)**	**Target (1977-1980)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
Finland	3665	2520	2581	0.795	0.805	0.579; 1.120
Denmark	3582	3097				

The third target set by Finland was for 1989. It was set in 1979. Figure 7 and Table 7 show the performance of Finland in realising this target, once more compared to Denmark. Denmark did have a quantified road safety target of its own in the last year of the after period (1989), but was nevertheless the most suitable comparison country that could be found.

Figure 7: Finland compared to Denmark before and after adoption of a quantified road safety target in Finland 1979-1989

The mean odds ratio for Denmark as compared to Finland in the before period was 0.984, with a variance of 0.011. As shown in Figure 7, the target for improving safety was not realised. The number of road accident fatalities increased, rather than going down. The target appears to have been strongly influenced by the very favourable trend during the last half of the 1970s. This trend did not continue into the 1980s.

Finland did nevertheless perform slightly better than Denmark during the years 1979-1989. The difference in safety performance was not statistically significant at the 5% level.

A new quantified road safety target was set in 1981, also referring to the year 1989. This target was even more ambitious than the target set in 1979, aiming for 300 fatalities in 1989. Figure 8 and Table 8 show the success in realising this target.

Table 7: Estimated effect of quantified road safety target in Finland 1979-1989

	Number of fatalities			Estimated effect		
Country	**Before (1967-1978)**	**After (1979-1989)**	**Target (1979-1989)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
Finland	11256	6591	5001	0.931	0.944	0.765; 1.164
Denmark	12159	7650				

Figure 8 looks pretty much like Figure 7. The favourable trend of the 1970s did not continue in the 1980s. Denmark matched Finland well in the before period (1971-1980), with a mean odds ratio of 0.991 and a variance of this odds ratio of 0.021.

Figure 8: Finland compared to Denmark before and after adoption of a quantified road safety target in Finland 1981-1989

Despite not realising the target, Finland did perform slightly better than Denmark. The difference in safety performance was not statistically significant at the 5% level.

Table 8: Estimated effect of quantified road safety target in Finland 1979-1989

	Number of fatalities			Estimated effect		
Country	**Before (1971-1980)**	**After (1981-1989)**	**Target (1981-1989)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
Finland	8484	5390	3704	0.919	0.936	0.702; 1.248
Denmark	9008	6230				

The fifth quantified road safety target set in Finland was for the year 1994. By this time, Denmark had adopted a quantified road safety target of its own, and was therefore no longer eligible as a comparison country. Switzerland was chosen as comparison country. Figure 9 and Table 9 present the evaluation of the target set for 1994 in Finland.

Figure 9: Finland compared to Switzerland before and after adoption of a quantified road safety target in Finland 1987-1994

Switzerland was not a perfect comparison for Finland. The mean odds ratio was 1.031, with a variance of 0.011. As shown in Figure 9, the number of fatalities did start to decline again in Finland after the increase during the 1980s. A similar, and indeed greater, decline occurred in Switzerland. Switzerland performed substantially better than Finland. The difference in safety performance between these two countries was statistically significant at the 5% level, to the disadvantage of Finland.

Does it make sense to compare countries as different as Finland and Switzerland in this respect? This is obviously a question that needs to be discussed carefully. This discussion is reserved for the next chapter. For now, suffice it to note that Switzerland did in fact perform better than Finland, despite the fact that it did not have a quantified road safety target in this period.

Table 9: Estimated effect of quantified road safety target in Finland 1987-1994

	Number of fatalities			Estimated effect		
Country	Before (1978-1986)	After (1987-1994)	Target (1987-1994)	Crude odds ratio	Adjusted odds ratio	95% CI
Finland	5233	4814	3807	1.371	1.340	1.085; 1.656
Switzerland	10031	6732				

The sixth and final target to be evaluated for Finland was set in 1989 for the year 2000. For the evaluation of this target, Japan was chosen as comparison. Figure 10 and Table 10 present the results of the evaluation.

Figure 10: Finland compared to Japan before and after adoption of a quantified road safety target in Finland 1989-2000

The choice of Japan as a comparison for Finland may seem even more odd than the choice of Switzerland. The two countries are located at opposite sides of the planet, almost as far away from each other as two countries can possibly get. Their cultures are different. Yet, their road safety history shows striking similarities. Both countries had a pretty bad road safety record in the early 1970s, but were able to improve this record very much by 1980. Both countries suffered setbacks in the 1980s, not being able to sustain the reductions in the number of fatalities they had accomplished in the previous decade. They matched very well during the years 1981-1990, with a mean odds ratio of 0.992 and a variance of this odds ratio of 0.005.

Table 10: Estimated effect of quantified road safety target in Finland 1989-2000

	Number of fatalities			Estimated effect		
Country	**Before (1981-1990)**	**After (1991-1999)**	**Target (1991-1999)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
Finland	6039	4307	4968	0.803	0.813	0.704; 0.939
Japan	126303	112116				

According to Table 10, Finland did perform better than Japan during the years 1991-1999. The difference in road safety performance was statistically significant at the 5% level.

The overall impression from the evaluations of the quantified road safety targets set in Finland is mixed. Some of the targets appear to have been effective, others do not appear to have been very successful. Possible explanations for these mixed results will be discussed in the next chapter.

5.1.4 France

The target set for France in 1997 has been assumed to be effective from the same year. Hence its effectiveness in the years 1997, 1998, and 1999 can be evaluated. Figure 11 and Table 11 present the results of this evaluation. Italy was chosen for comparison. It matches France quite well (mean odds ratio 0.986; variance of odds ratio 0.001).

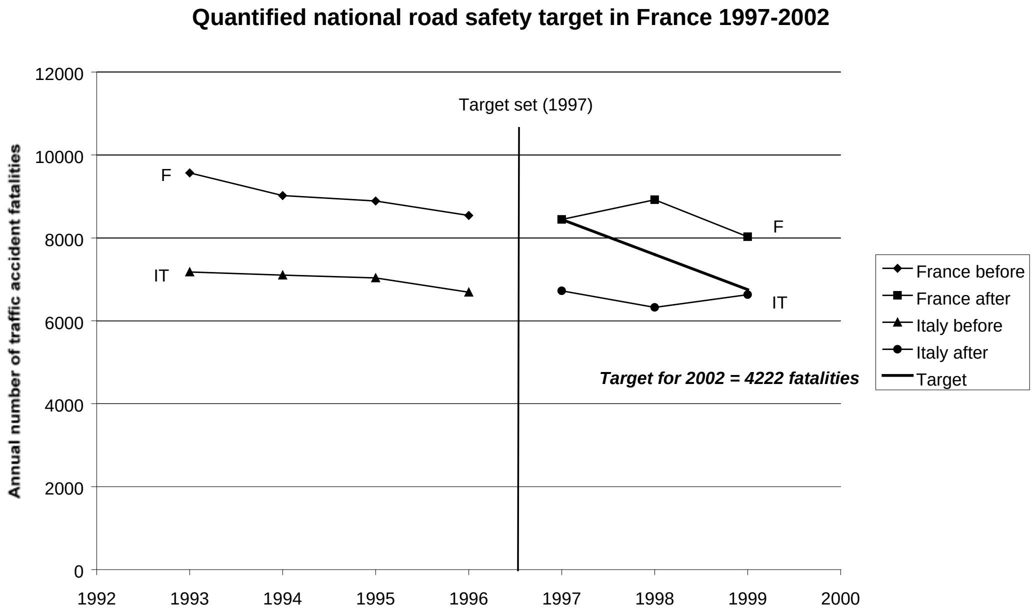

Figure 11: France compared to Italy before and after adoption of a quantified road safety target in France 1997-2002

The target set in France is very demanding, aiming for a 50% reduction of the number of fatalities in just five years. This corresponds to an average annual reduction of 12.9%.

Table 11: Estimated effect of quantified road safety target in France 1997-2002

	Number of fatalities			Estimated effect		
Country	Before (1993-1996)	After (1997-1999)	Target (1997-1999)	Crude odds ratio	Adjusted odds ratio	95% CI
France	36019	25391	22799	1.003	1.018	0.953; 1.088
Italy	28002	19683				

So far, the target set in France does not appear to have been a success. Road safety has not improved more in France than in Italy, nor has the annual rate of reduction of the number of fatalities in France become greater after the target was adopted than it was before.

Previous research (Elvik 1993) has concluded that highly ambitious targets are more effective than less ambitious targets. But perhaps there is a point beyond which an ambitious target is no longer taken seriously, because it may appear hopeless to realise. Reducing the number of fatalities by 50% in five years is very ambitious. But there are examples of governments that have been able to do so, notably the government of the state of Victoria in Australia in the early part of the

1990s. But the population of the state of Victoria is much smaller than the population of France. Although France has a tradition of a strong central government, governing France is likely to be much more complex than governing an Australian state.

5.1.5 Hungary

Hungary has taken a lead among the former East Bloc states in pursuing an ambitious road safety programme after the collapse of Communism in the late 1980s. It has made the use of daytime running lights mandatory. Speed limits in urban areas have been reduced from 60 to 50 km/h. A quantified road safety target for 2000 was set in 1993. Figure 12 and Table 12 present the results of an evaluation of the effectiveness of this target.

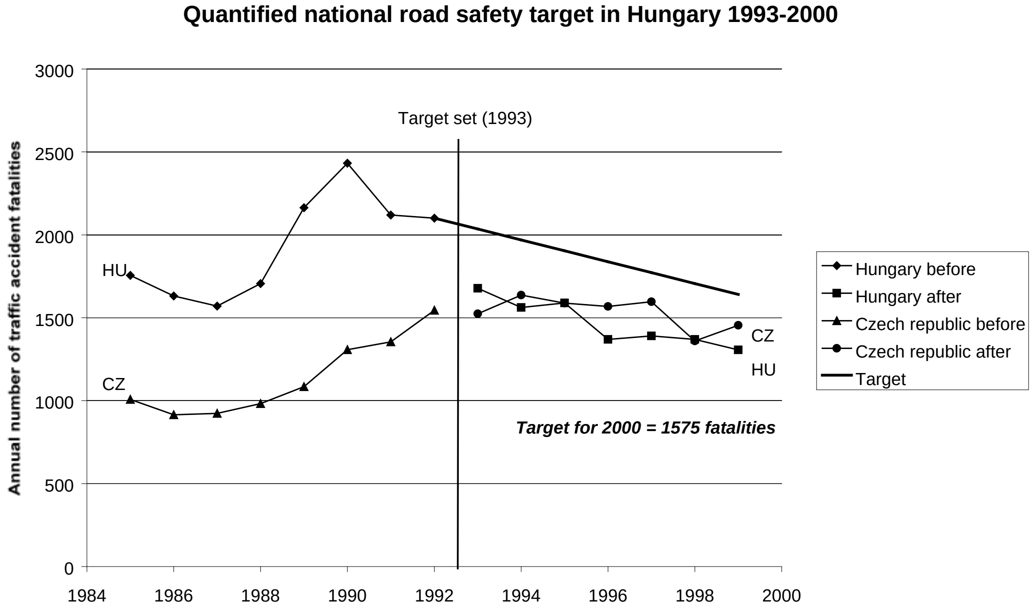

Figure 12: Hungary compared to the Czech Republic before and after adoption of a quantified road safety target in Hungary 1993-2000

The Czech Republic was chosen as a comparison for Hungary. These countries matched quite well in the before period (mean odds ratio 0.970; variance of odds ratio 0.011). As is evident from Figure 12, there was a sharp increase in the number of road accident fatalities in both countries during the years when Communism collapsed. Hungary was able to turn this trend in 1991, whereas the Czech Republic has been less successful.

Table 12: Estimated effect of quantified road safety target in Hungary 1993-2000

	Number of fatalities			Estimated effect		
Country	Before (1985-1992)	After (1993-1999)	Target (1993-1999)	Crude odds ratio	Adjusted odds ratio	95% CI
Hungary	15482	10266	12866	0.564	0.589	0.478; 0.726
Czech Republic	9120	10729				

The quantified road safety target in Hungary is associated with a substantially better road safety performance than in the Czech Republic. The number of fatalities has been reduced by more than 40% in Hungary, compared to the Czech Republic. This difference in safety performance is statistically significant at the 5% level. The target set in Hungary has been fully realised. Indeed, this target was realised during the first year it was effective.

5.1.6 Iceland

Iceland is a tiny country on the outskirts of Europe. It has long been known for its successful road safety record. Its population fatality rate is one of the lowest in the World for a highly motorised country. Iceland has recently adopted a quantified road safety target. This target applies to the total number of injuries in road accidents, not just to fatalities (Traffic Safety Committee 1997). To maintain comparability with other countries, the target has been evaluated in terms of fatalities only. Figure 13 and Table 13 present the results of the evaluation.

Figure 13: Iceland compared to Luxembourg before and after adoption of a quantified road safety target in Iceland 1997-2000

Finding a suitable comparison for Iceland was not easy. Iceland generally prefers to compare itself to the other Nordic countries, but most of these had a quantified road safety target during either the before or after period. Moreover, the criterion for successful matching has to be interpreted somewhat liberally in this case, since there are large random fluctuations from year to year in the number of fatalities in Iceland (the counts in the before series are 17, 12, 24, and 10). There is no point in trying to match such random fluctuations. Luxembourg was chosen for comparison, being itself a small country. It did not match Iceland perfectly on a year to year basis, but the overall matching was adequate (mean odds ratio 1.036; variance of odds ratio 0.581).

Table 13 indicates that Iceland has not succeeded in reducing the number of fatalities after it set the quantified road safety target. In fact, Luxembourg has been much more successful than Iceland.

Table 13: Estimated effect of quantified road safety target in Iceland 1997-2000

	Number of fatalities			Estimated effect		
Country	**Before (1993-1996)**	**After (1997-1999)**	**Target (1997-1999)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
Iceland	63	63	51	1.617	1.828	0.390; 8.577
Luxembourg	283	175				

There has been a net increase of more than 80% in the number of fatalities in Iceland from before to after a quantified target was set, compared to Luxembourg. Random fluctuations may in part account for this. The difference in safety performance is not statistically significant and associated with a very wide confidence interval. It would seem that Iceland has had a run of bad luck in the years after it set the safety target.

5.1.7 The Netherlands

The Netherlands has set two quantified road safety targets: One for the year 2000, the other for the year 2010. Both targets were in force during the years before 2000. The effects of these targets on the safety performance of the Netherlands has been evaluated by comparing the Netherlands to Belgium. Figure 14 and Table 14 present the results with respect to the target set for 2000.

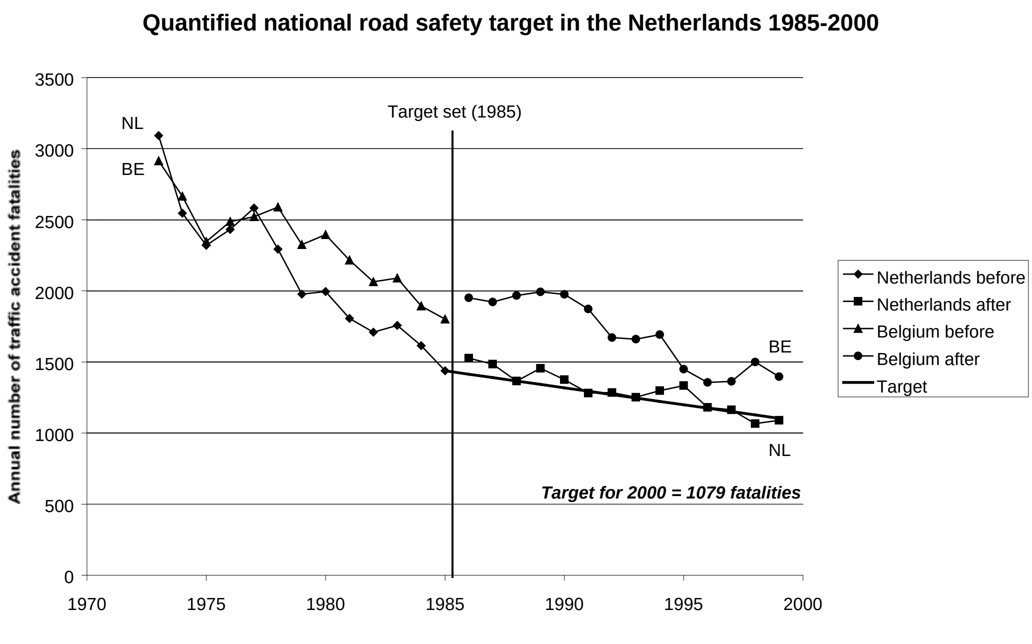

Figure 14: The Netherlands compared to Belgium before and after adoption of a quantified road safety target in the Netherlands 1985-2000

Belgium matched the Netherlands quite well in the before period (mean odds ratio 0.978; variance of odds ratio 0.003).

The target set for 2000 is associated with an improved road safety performance in the Netherlands, amounting to a net improvement of 14% compared to Belgium. This difference is statistically significant at the 5% level of significance.

Table 14: Estimated effect of quantified road safety target in The Netherlands 1985-2000

	Number of fatalities			Estimated effect		
Country	**Before (1973-1985)**	**After (1986-1999)**	**Target (1986-1999)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
The Netherlands	27567	18159	17623	0.840	0.859	0.784; 0.941
Belgium	30311	23772				

Turning to the target set for 2010, and once more using Belgium for comparison Figure 15 and Table 15 present the results of an evaluation. Belgium matched the Netherlands well in the before period (mean odds ratio 0.980; variance of odds ratio 0.002).

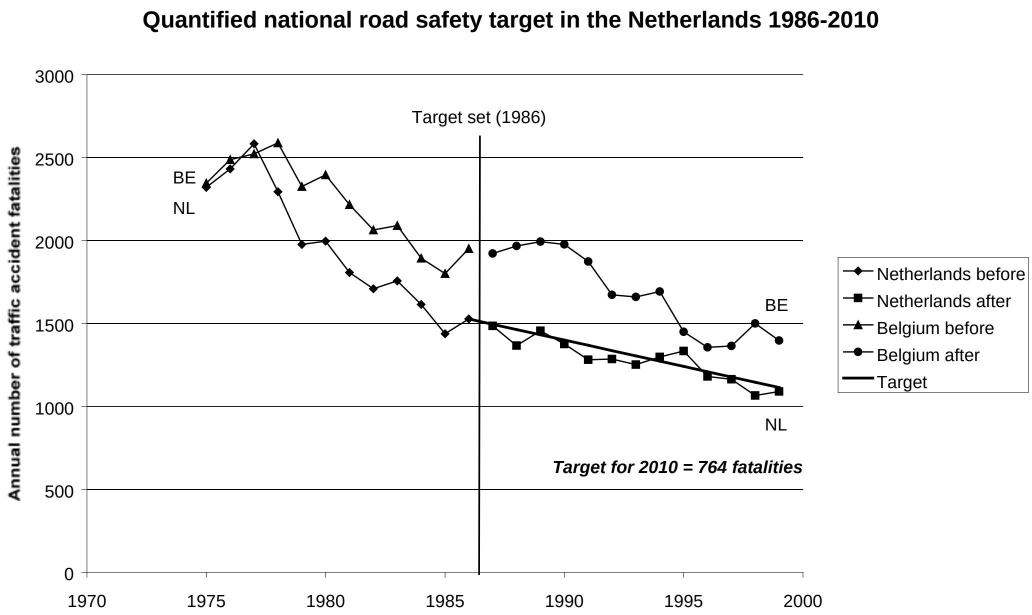

Figure 15: The Netherlands compared to Belgium before and after adoption of a quantified road safety target in the Netherlands 1986-2010

The long term target set for the Netherlands is associated with an improved road safety performance. The net difference, compared to Belgium, is an improvement of 11% during the years 1987-1999. This difference is statistically significant at the 5% level.

Table 15: Estimated effect of quantified road safety target in The Netherlands 1986-2010

	Number of fatalities			Estimated effect		
Country	**Before (1976-1986)**	**After (1987-1999)**	**Target (1987-1999)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
The Netherlands	23456	16632	16957	0.867	0.890	0.797; 0.994
Belgium	26682	21821				

5.1.8 New Zealand

New Zealand has had a quantified road safety target since 1990. The first target set was a short term target for the year 1994. The next target was for 2001. Recently, a new target has been proposed for 2010. This target is, however, too new to be evaluated at this time. The evaluation therefore comprises the targets for 1994 and 2001.

Figure 16 and Table 16 present the results of the evaluation of the target set for 1994. Canada was chosen for comparison. It matched the accident experience of New Zealand quite well during the before period (mean odds ratio 0.995; variance 0.002).

Quantified national road safety target in New Zealand 1990-1994

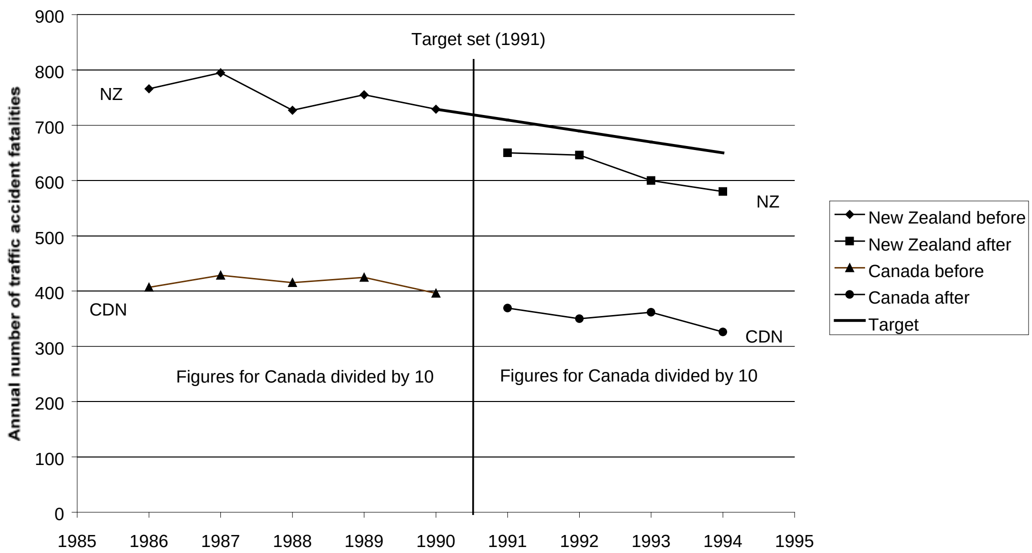

Figure 16: New Zealand compared to Canada before and after adoption of a quantified road safety target in New Zealand 1991-1994

The target set for 1994 in New Zealand was realised. New Zealand did have a slightly better safety performance during this period than Canada, but the difference was not statistically significant at the 5% level.

Table 16: Estimated effect of quantified road safety target in New Zealand 1991-1994

	Number of fatalities			Estimated effect		
Country	**Before (1986-1990)**	**After (1991-1994)**	**Target (1991-1994)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
New Zealand	3772	2476	2719	0.967	0.973	0.877; 1.079
Canada	20714	14067				

For the evaluation of the target set in New Zealand for the year 2001, Japan was chosen as the comparison country. Japan matched New Zealand very well (mean odds ratio 1.008; variance of odds ratio 0.008). Figure 17 and Table 17 present the results of the evaluation.

Figure 17: New Zealand compared to Japan before and after adoption of a quantified road safety target in New Zealand 1990-2001

According to Table 17, New Zealand did improve its road safety performance after the target for 2001 was set. Compared to Japan, the net improvement was 18.5% during the years 1990-1999. The difference in safety performance is statistically significant at the 5% level.

Table 17: Estimated effect of quantified road safety target in New Zealand 1991-1994

	Number of fatalities			Estimated effect		
Country	**Before (1979-1989)**	**After (1990-1999)**	**Target (1990-1999)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
New Zealand	7598	5850	6026	0.815	0.815	0.682; 0.975
Japan	134100	126710				

5.1.9 Norway

Norway has long had one of the best road safety records of the highly motorised countries. Road safety has been greatly improved since 1970, which was the worst year on record, with 560 people killed in road accidents. There have, however, been setbacks along the way. One of the longest of these occurred in the years 1981-1986, when there was an increase in the number of road accident fatalities. The year 1986 was the worst since 1976, and this persuaded Norwegian politicians to accept the idea of setting a quantified national road safety target for the first time. The target set was, however, not a very ambitious one. It was simply to prevent the number of road accident fatalities from increasing. The target was set for 1993, and was based on the mean annual number of fatalities during the years 1984-1986, which was 420. Figure 18 and Table 18 present an evaluation of the success of this road safety target.

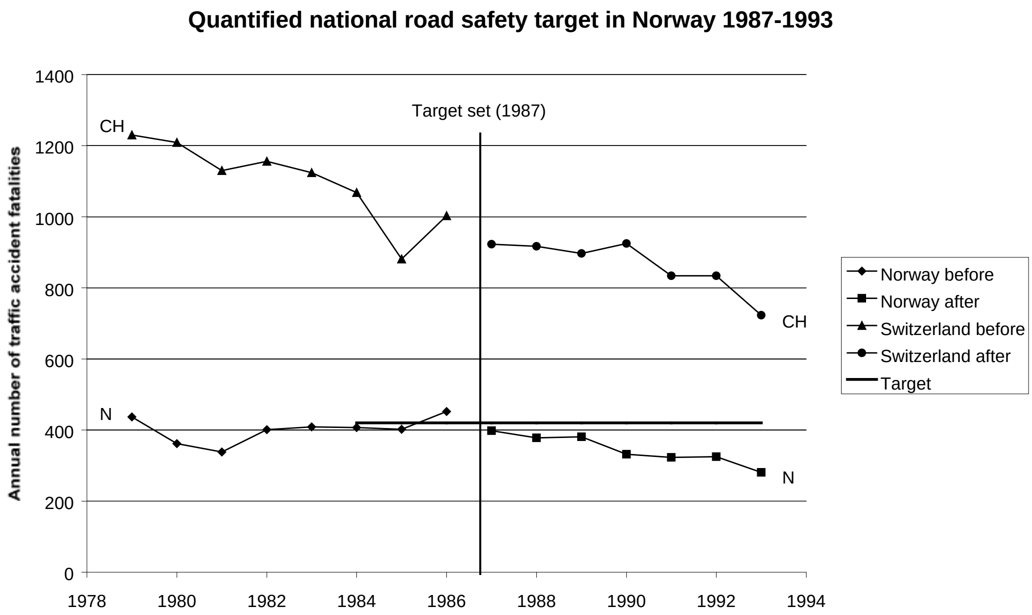

Figure 18: Norway compared to Switzerland before and after adoption of a quantified road safety target in Norway 1987-1993

Finding a suitable comparison country for Norway was not easy. All its Nordic neighbouring countries had quantified road safety targets sometime during the before or after period, and were therefore not eligible for comparison. Switzerland was chosen for comparison, although it did not match Norway as closely as one would ideally like (mean odds ratio = 1.040; variance of odds ratio = 0.007).

Table 18 indicates that the road safety target set for Norway did not result in a better safety performance, compared to the performance of Switzerland in the same period. Although the number of road accident fatalities in Norway fell substantially in the period when this target was effective, an even greater reduction took place in Switzerland. The difference between the two countries was, however, not statistically significant at the conventional 5% level.

Table 18: Estimated effect of quantified road safety target in Norway 1987-1993

	Number of fatalities			**Estimated effect**		
Country	**Before (1979-1986)**	**After (1987-1993)**	**Target (1987-1993)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
Norway	3208	2418	2940	1.096	1.062	0.891; 1.266
Switzerland	8801	6053				

There are several features of this result that deserve to be discussed more in detail. The full discussion will be the subject of chapter 6. However, a couple of points will be made in passing.

First, the Norwegian experience illustrates in a very striking way the pitfalls in extrapolation of recent accident trends as a basis for setting road safety targets. The years immediately before the target was set were pretty bad, especially 1986. Traffic was increasing rapidly. Figure 19 shows an extrapolation of the trend from 1980 to 1986 into the after period.

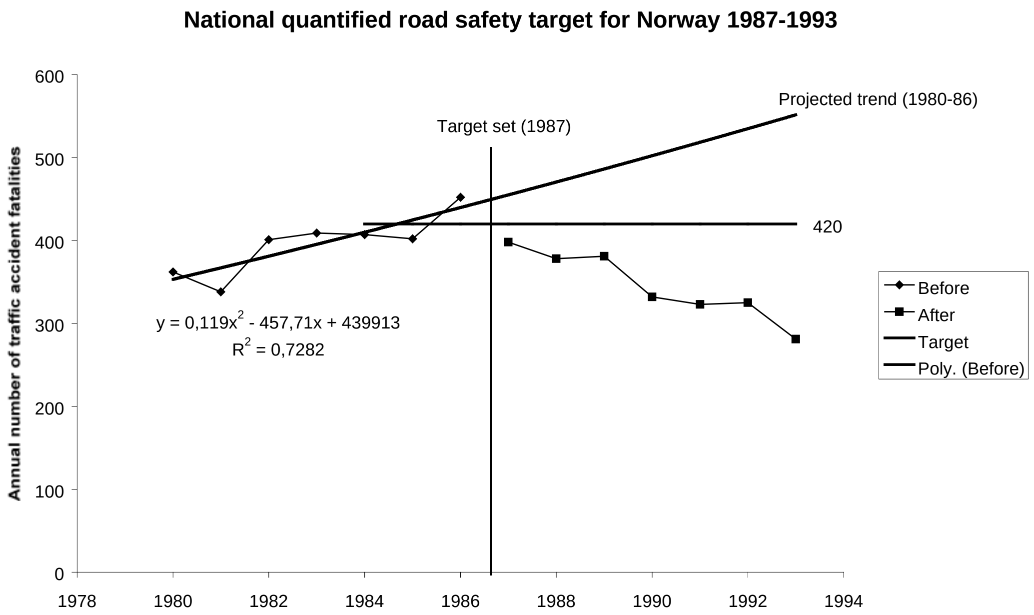

Figure 19: Projection of trend in road accident fatalities 1980-1986 in Norway to the years 1987-1993

The trend fits the before data quite well (R-squared = 0.7282). As a prediction for after period, on the other hand, it failed completely. With the benefit of hindsight, the target that was set looks extremely conservative. Such an assessment is, of course, nothing but hindsight bias. At the time the target was set, the trend was very worrying indeed. Nobody could have predicted that this trend would turn drastically just after the target had been set. History does not go on repeating itself for ever, and can be a treacherous foundation for making predictions.

The second point to be made is that by any reasonable standard, the target that was set would be regarded as a success. In fact, things went even better than the target aimed for. In all years after the target was set, the number of road accident fatalities was below the targeted number. This fact is probably not attributable to road safety policy exclusively. Nevertheless it resembles a paradox when a formal evaluation returns exactly the opposite verdict: This target was not at all a success, and road safety performance in Norway was actually a little worse than in Switzerland, which did not have a quantified road safety target.

The resemblance of a paradox is reinforced when the results of the evaluation study are assessed in terms of the three observations consistent with the hypothesis that a quantified road safety target improves safety performance, that were deduced in chapter 4. To repeat, these three criteria were:

1. The mean annual percentage change of the number of road accident fatalities should shift towards a greater reduction in the after period than in the before period.
2. The mean annual percentage change of the number of road accident fatalities in the after period should be more favourable in the target country than in the comparison country.

3. The net change in the long term trend for the mean annual number of road accident fatalities from the before to the after period should favour the target country.

Let us examine how the evaluation of the quantified road safety target set in Norway fares according to these criteria. In the before period, the mean annual change in the number of road accident fatalities in Norway was +1.1% per year. In the after period, this switched to –5.5%. Hence criterion 1 is fulfilled.

In Switzerland, the mean annual change in the number of road accident fatalities in the after period was –3.8%. Criterion 2 is therefore fulfilled: The annual rate of decline was greater in Norway than it was in Switzerland.

From the before period to the after period, the annual rate of change in the number of road accident fatalities in Norway changed from +1.1% to –5.5%, a net change of 6.6%. The corresponding change in Switzerland was from –2.5% to –3.8%, a net change of 1.3%. Hence, the net change in trend favours Norway.

All these observations are consistent with the hypothesis that the target set in Norway was effective. Yet, the evaluation study reached a different conclusion. Is there a statistical paradox here? To explore this question, the evaluation was repeated by means of standardised numbers.

Let both before series start at 1000. After one year, the number has grown to 1011 in Norway, assuming a mean annual growth rate of 1.1%. After two years, the number has grown to 1022, and so on. Summing for eight years (the length of the before period), the value is 8315. For Switzerland, the corresponding sum is 7334. Let both after series start at 1000 as well. The sum over seven years then becomes 5945 for Norway and 6251 for Switzerland. An odds ratio based on these numbers is 0.951, indicating a net gain in safety performance of about 5% for Norway.

Clearly some sort of statistical paradox leads to the opposite result in the evaluation study. The nature of this paradox will be examined more closely in Chapter 6.

5.1.10 Poland

The GAMBIT road safety programme for Poland, published in abbreviated version by the Technical University of Gdansk in 1996 contains a quantified road safety target of not more than 6000 fatalities in 2001. It is not altogether clear from which date this target has been effective, but in the analysis presented here, it assumed that the target has been in force from 1996.

Poland is not a member of the IRTAD database, which is used as the main source of data in this study. Official Danish accident statistics contains main figures from road accidents statistics for a large number of countries, and was used the source of data for Poland. Finding a suitable comparison country was difficult. Portugal was chosen. It matched Poland quite well (mean odds ratio = 1.011; variance of odds ratio = 0.011). Figure 20 and Table 19 present the results of an evaluation of the quantified road safety target for Poland, using Portugal as comparison.

Figure 20: Poland compared to Portugal before and after adoption of a quantified road safety target in Poland 1996-2001

According to the results presented in Table 19, setting a quantified road safety target in Poland did not improve road safety performance, compared to Portugal. There was a net increase of 10% in the number of fatalities. This increase was not statistically significant at the 5% level.

Table 19: Estimated effect of quantified road safety target in Poland 1996-2001

	Number of fatalities			Estimated effect		
Country	**Before (1991-1995)**	**After (1996-1998)**	**Target (1996-1998)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
Poland	34832	20749	19800	1.105	1.104	0.896; 1.360
Portugal	10953	5904				

5.1.11 Spain

Spain set a quantified road safety target for 1999 in 1992. The road safety experience in Spain before and after this target was set has been compared to that of Portugal. Portugal matched Spain very well (mean odds ratio 0.999; variance of odds ratio 0.009). Figure 21 and Table 20 present the results of the study evaluating the effectiveness of the quantified road safety target set in Spain.

Spain did improve its road safety performance after the target was set. The net improvement in the period 1993-1999 amounts to 14%, compared to the experience of Portugal in the same period. The difference is not statistically significant at the 5% level, but comes close to being so.

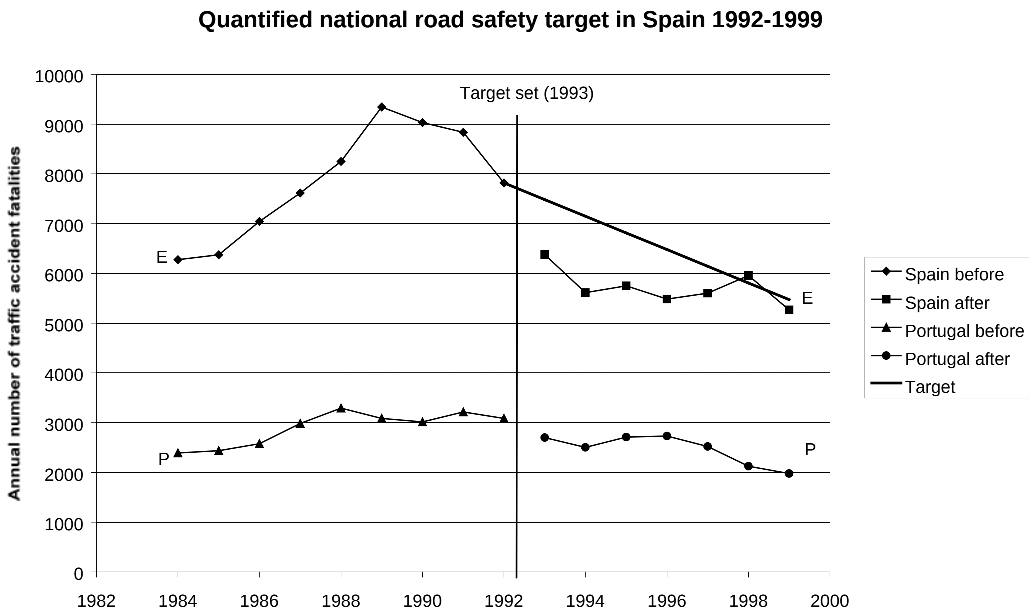

Figure 21: Spain compared to Portugal before and after adoption of a quantified road safety target in Spain 1992-1999

Table 20: Estimated effect of quantified road safety target in Spain 1992-1999

	Number of fatalities			Estimated effect		
Country	Before (1984-1992)	After (1993-1999)	Target (1993-1999)	Crude odds ratio	Adjusted odds ratio	95% CI
Spain	70591	40056	45346	0.857	0.862	0.715; 1.039
Portugal	26091	17270				

5.1.12 Sweden

Sweden has always had ambitious programmes to improve road safety, and has succeeded better in doing so than many other countries. It ranks among the safest of the highly motorised countries. Two quantified road safety targets set by the Swedish government are evaluated in this report. One target was set for 2000, another for 2007.

Finding a suitable comparison for Sweden was difficult. The United States was used initially, but it has recently adopted a quantified road safety target and is therefore not eligible. Japan was finally used, although it did not match the accident trend for Sweden in the before period as closely as one would ideally like (mean odds ratio = 0.974; variance of odds ratio = 0.004). Other countries, such as Finland, might have matched Sweden more closely, but had a quantified road safety target of their own.

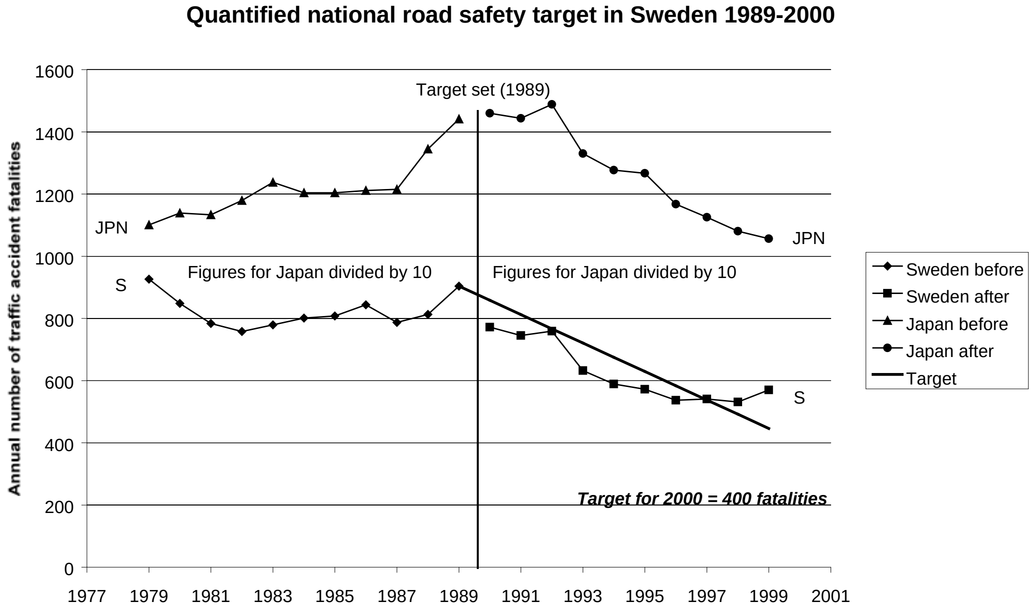

Figure 22: Sweden compared to Japan before and after adoption of a quantified road safety target in Sweden 1989-2000

Table 21 shows that Sweden has performed better than Japan in improving road safety in the decade 1990-1999. The net difference is 25%. This difference is statistically significant at the 5% level of significance.

Table 21: Estimated effect of quantified road safety target in Sweden 1989-2000

	Number of fatalities			Estimated effect		
Country	**Before (1979-1989)**	**After (1990-1999)**	**Target (1990-1999)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
Sweden	9052	6248	6521	0.729	0.752	0.661; 0.854
Japan	134102	126923				

In 1996, Sweden set a new target for 2007, of reducing the number of road accident fatalities by 50%, down to 270 in 2007. Once again, finding a suitable comparison country in order to evaluate the effects of this target was difficult. Canada was chosen, and matches Sweden adequately (mean odds ratio in before period = 0.987; variance of odds ratio = 0.013).

Figure 23 and Table 22 present the results of the evaluation. The data indicate that the target set for 2007 has so far been less successful than the target set for 2000. There has been a net reduction of the number of fatalities in Sweden compared to Canada, but the difference is very small and far from statistically significant at the 5% level.

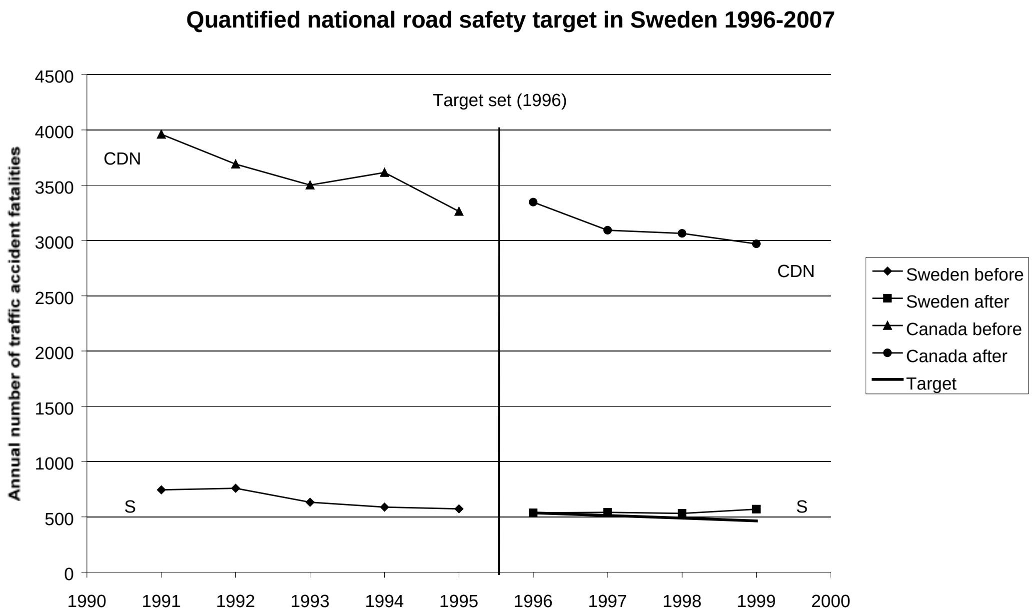

Figure 23: Sweden compared to Canada before and after adoption of a quantified road safety target in Sweden 1996-2007

Table 22: Estimated effect of quantified road safety target in Sweden 1996-2007

	Number of fatalities			Estimated effect		
Country	**Before (1991-1995)**	**After (1996-1999)**	**Target (1996-1999)**	**Crude odds ratio**	**Adjusted odds ratio**	**95% CI**
Sweden	3297	2179	2002	0.955	0.973	0.772; 1.226
Canada	18030	12472				

5.1.13 United Kingdom

The United Kingdom adopted a quantified road safety target in 1986, based on the mean annual number of road accident fatalities in 1981-1985. The target applied to the year 2000. France has been used as a comparison country. During most the period covered by the study, France did not have a quantified road safety target. France matched the United Kingdom quite well during the before period (mean odds ratio = 1.010; variance of odds ratio = 0.007).

Figure 24 and Table 23 present the results of the study. The data presented in the Figure and the Table show that the quantified target set in the United Kingdom is associated with a slight improvement in safety performance compared to France. The difference is around 5% and is not near statistical significance at the conventional 5% level.

Figure 24: The United Kingdom compared to France before and after adoption of a quantified road safety target in the United Kingdom 1986-2000

Table 23: Estimated effect of quantified road safety target in the United Kingdom 1986-2000

	Number of fatalities			Estimated effect		
Country	Before (1973-1985)	After (1986-1999)	Target (1986-1999)	Crude odds ratio	Adjusted odds ratio	95% CI
United Kingdom	84428	62146	67589	0.955	0.950	0.806; 1.119
France	179970	138670				

5.1.14 The United States of America

The United States of America has set a national road safety target of no more than 33500 road accident fatalities in 2008. It will be assumed that this target has been effective in the years 1997, 1998, and 1999. The United States is unique among the motorised countries. Unlike almost all other countries, it experienced an increase in the number of road accident fatalities from 1992 to 1996. At this time, many other motorised countries had set quantified road safety targets. A suitable comparison country for the United States could not be found. The analysis is therefore based on fatality data for the United States exclusively.

Figure 25 presents these data. Figure 25 clearly shows that there was an increase in the number of road accident fatalities in the before period, and that there has been a decline in the after period. So far, however, the number of fatalities remains higher than the targeted number. Figure 25 resembles Figure 19 presented for Norway, in discussing the dangers of extrapolating a trend. The same point of view applies to the United States. There are many ways of analysing the data in Figure 25.

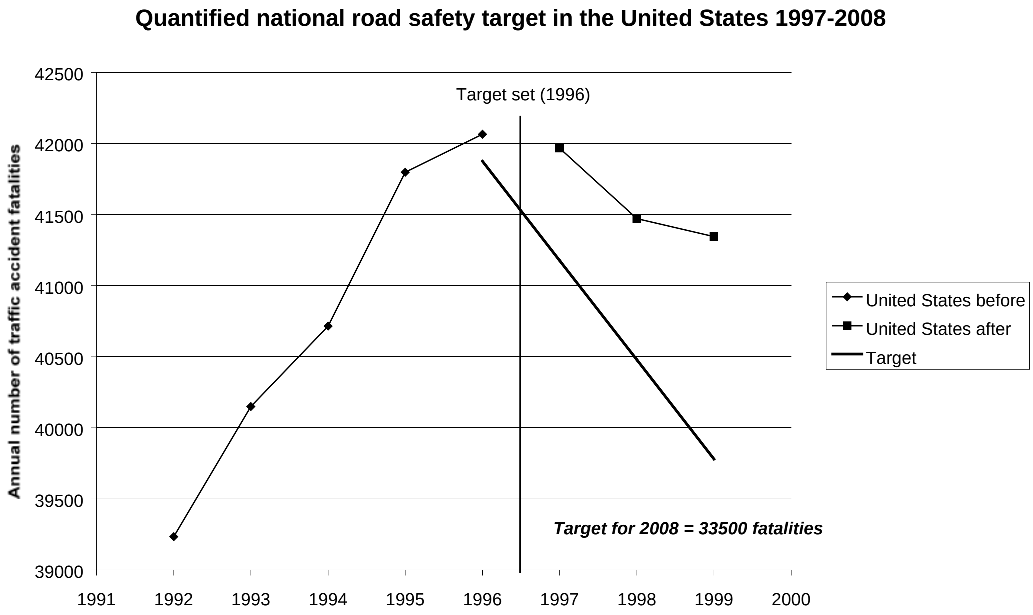

Figure 25: Annual number of road accident fatalities in the United States before and after adoption of a quantified road safety target

The mean annual number of fatalities in the before period (1992-1996) was 40793. The mean annual number of fatalities in the after period (1997-1999) was 41594. These numbers do not indicate that the quantified road safety target has been effective so far, showing an increase of 2%. This increase is not statistically significant at the 5% level (0.957; 1.087). The confidence interval was estimated by treating the variance around the mean annual number of fatalities in the before period as analogous to the variance of the odds ratio in the comparison group method.

5.2 Before-and-after study of local road safety targets

This section presents the results of the before-and-after studies that have been made of local road safety targets. Three sets of such targets have been identified:

1. Targets set by Norwegian counties in the 1980s, and previously analysed by Elvik (1993).
2. Targets set by Australian states, identified in an OECD report on targeted road safety programmes (OECD 1994).
3. Targets set by local governments in Germany, in particular the city of Darmstadt (Schlabbach 1990).

5.2.1 Local road safety targets in Norway

Table 24 presents the results of a study evaluating the effects of local road safety targets set by Norwegian counties.

Table 24: Evaluation of quantified road safety targets set by Norwegian counties

	Number of police reported injury accidents					
	Target counties		Comparison counties		Measure of effect	
Group of targets	**Before**	**After**	**Before**	**After**	**Odds ratio**	**95% CI**
1 (-50%, 10 yrs)	1868	1487	4454	4584	0.773	0.714; 0.838
2 (-26%; 6 yrs)	1928	1943	4454	4584	0.979	0.908; 1.056
3 (-18%; 5 yrs)	1926	1950	4454	4584	0.984	0.912; 1.061
4 (-31%; 8 yrs)	1713	1686	780	865	0.888	0.789; 0.999
5 (-12%; 7 yrs)	2010	2181	780	865	0.978	0.873; 1.097
6 (-0%; 4 yrs)	1068	1058	780	865	0.893	0.785; 1.016
7 (-27%; 7 yrs)	1835	1895	780	865	0.931	0.829; 1.046
8 (0%; 8 yrs)	1006	1227	780	865	1.100	0.968; 1.250

According to Table 24, most of the targets set by Norwegian counties were associated with modest gains in road safety, in the order of 5-10%. Most of these changes were not statistically significant. There is a tendency for more ambitious targets to be more successful than less ambitious targets. For a further discussion, see Elvik (1993).

5.2.2 Local road safety targets in Australia

The targets that have been identified were set by New South Wales, Victoria, Western Australia, and the Northern Territories. All targets had 1991 as the base year and 2000 as the target year. Table 25 presents the results of an evaluation study.

Table 25: Evaluation of quantified road safety targets set by Australian states

	Number of road accident fatalities					
	Target states		Comparison states		Measure of effect	
State	**Before**	**After**	**Before**	**After**	**Odds ratio**	**95% CI**
New South Wales	663	608	673	545	1.132	0.967; 1.326
Victoria	503	407	673	545	0.999	0.841; 1.188
Western Australia	207	213	673	545	1.271	1.018; 1.587
Northern Territory	67	52	673	545	0.958	0.656; 1.401

Australian states that have set quantified road safety targets did not consistently perform better in improving road safety than other Australian states during the period 1991-2000. In fact, one state did significantly worse than the other states. The results presented in Table 25 do not lend support to the hypothesis that quantified road safety targets are effective in improving road safety performance.

5.2.3 Local road safety targets in Germany

According to Schlabbach (1990), the city of Darmstadt set a target of reducing the number of injury accidents by 10% in one year. He presents a table showing that the number of injury accidents in Darmstadt went down from 1152 during the year from October 1987 through September 1988 to 1030 during the year from October 1988 through September 1989. No comparison data are provided. Data for Germany as a whole show, however, that there was hardly any change in the number of injury accident accidents in built-up areas during this period. The recorded number was 224,697 in 1987/88 and 224,960 in 1988/89. The net accident reduction in Darmstadt was 10.7% (odds ratio = 0.893), which is statistically significant at the 5% level (0.821; 0.972).

5.3 Synthesis of results of before-and-after studies

Does it make sense to combine the results of the before-and-after studies in order to test the hypotheses proposed in Chapter 2? In order to answer this question, a simple diagnostic analysis of the consistency and distribution of results has been made.

5.3.1 Testing the consistency and homogeneity of the results of the before-and-after studies

Figure 26 presents a funnel plot of the results of the before-and-after studies. The horizontal axis measures the value of each result, the vertical axis gives the statistical weight assigned to it, according to a fixed effects model of meta-analysis.

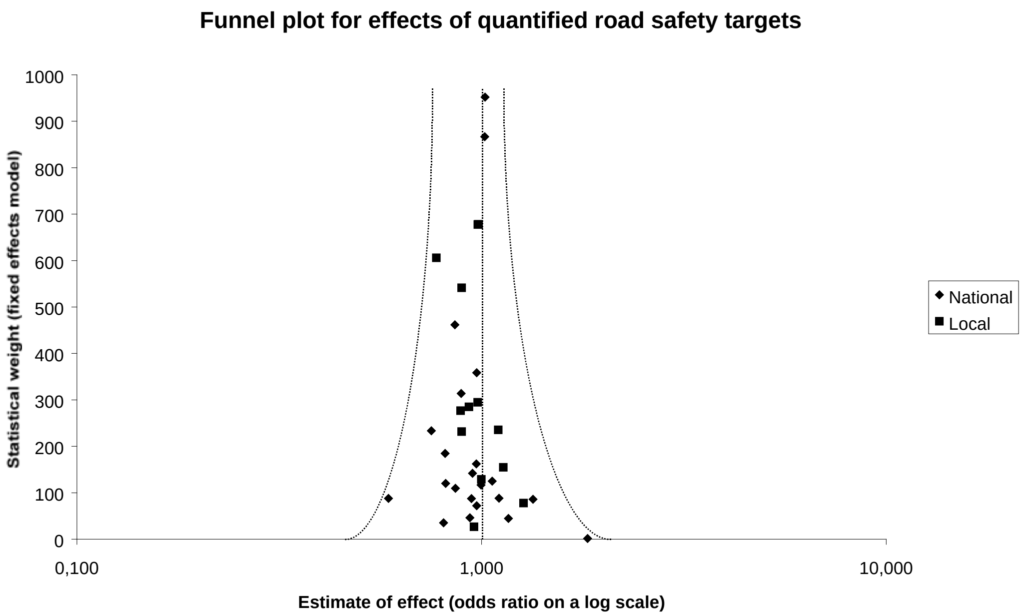

Figure 26: Funnel plot of results of before-and-after studies of quantified road safety targets

The contours of a funnel, upside down, have been indicated around the data points. The results are fairly closely concentrated. There are 35 estimates of effect in total. 22 of these refer to national targets, 13 refer to local targets. 15 of the 22 estimates of the effect of national targets indicate a reduction of the number of road accident fatalities, 7 indicate an increase. 10 of the 13 estimates of the effect of local targets indicate that safety has been improved, 3 indicate that it has deteriorated.

The effects attributed to quantified road safety targets are generally quite small. 10 estimates for national targets are within plus or minus 10% from zero. 7 estimates for local targets are within this range. The funnel plot is symmetrical and unimodal. Figure 27, derived from the funnel plot, shows the distribution of estimates of effect by magnitude.

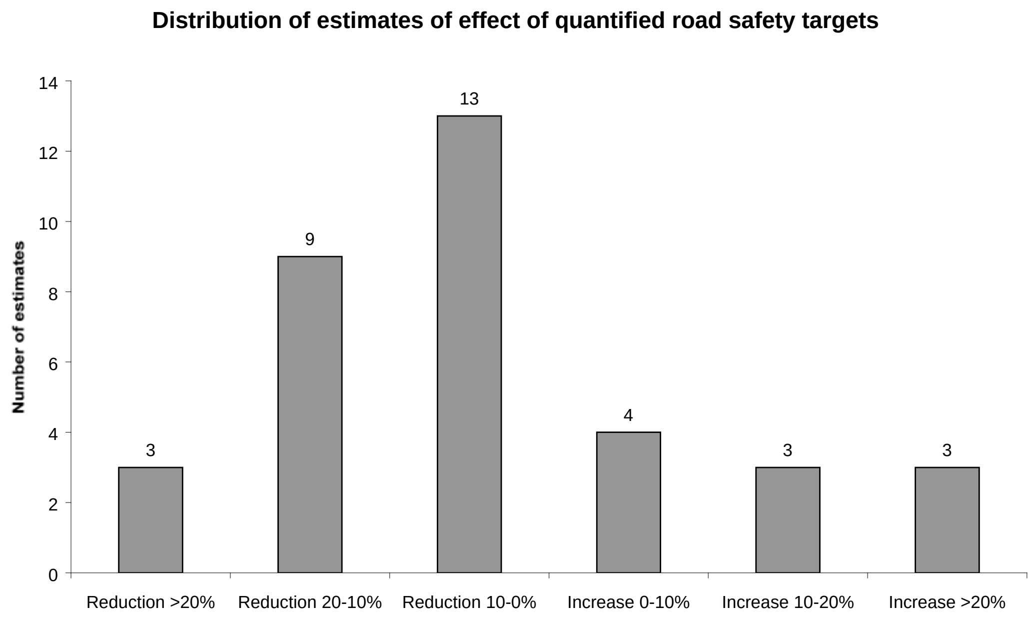

Figure 27: Distribution of estimates of effect of quantified road safety targets by magnitude

The distribution is slightly skewed towards a reduction of the number of fatalities or accidents, suggesting that the weighted mean effect of quantified road safety targets may be a small improvement in road safety. Both Figure 26 and Figure 27 are sufficiently well behaved in terms of modality (unimodal), dispersion (fairly narrow) and skewness (slightly skewed) for a formal synthesis of findings by means of a meta-analysis to make sense.

The homogeneity statistic for the 35 estimates of effect is highly significant (Chi-square 106.136, df = 34, P < 0.0000). Meta-analysis therefore relied on a random effects model. A fixed effects model was not fitted.

5.3.2 Meta-analysis of effects of quantified road safety targets

Results of the analysis using a random effects model, stated in terms of the weighted mean odds ratio associated with quantified road safety targets, are presented in Table 26. These results are based on the complete data set.

Hypothesis 1 is supported. Having a quantified road safety target is associated with a better road safety performance than not having such a target. The difference amounts to nearly 6% (mean odds ratio = 0.944) and is statistically significant at the 5% level.

Table 26: Weighted mean effects of quantified road safety targets. Random effects model. Complete data set

			Effects attributed to quantified road safety targets	
Hypothesis tested	**Values of variable**	**Number of targets**	**Mean odds ratio**	**95% CI**
1: Target vs no target	Target set	35	0.944	0.902; 0.988
2: Level of ambition	>-5%/year	12	0.911	0.840, 0.989
	-2-5%/year	16	0.950	0.888; 1.016
	<-2%/year	7	0.977	0.890; 1.072
3: Long vs short term	10- years	8	0.863	0.787; 0.946
	1-9 years	27	0.972	0.922; 1.023
4: National vs local	National	22	0.931	0.877; 0.988
	Local	13	0.963	0.898; 1.033
5: First vs revised	First target	24	0.954	0.904; 1.006
	Later targets	11	0.922	0.848; 1.002
6: High vs low risk	High risk	3	0.806	0.690; 0.940
	Medium risk	11	0.956	0.877; 1.043
	Low risk	8	0.953	0.865; 1.049
	Local targets	13	0.963	0.898; 1.033

Odds ratio 1.000 = no effect, <1.000 = reduced number of fatalities, >1.000 = increased number of fatalities

Hypothesis 2, regarding the level of ambition of a quantified road safety target, is also supported. Targets were classified in three groups with respect to the level of ambition:

1. Targets aiming for an annual reduction of more than 5% in the number of road accidents or the number of road accident fatalities,
2. Targets aiming for an annual reduction of between 2% and 5% in the number of road accidents or fatalities,
3. Targets aiming for an annual reduction of less than 2% in the number of road accidents or fatalities.

The largest difference in safety performance is associated with the most ambitious targets, the smallest difference is associated with the least ambitious targets.

Hypothesis 3 refers to the time horizon of a target. A distinction has been made between long-term targets, extending for 10 years or more, and short-term targets, extending for less than 10 years. The hypothesis was that long term targets are more effective than short term targets. This hypothesis is supported by the results presented in table 26.

Targets set by national governments are associated with a better safety performance than targets set by local governments. This finding supports hypothesis 4.

First generation road safety targets are not associated with a better road safety performance than later generations of road safety targets. This finding contradicts hypothesis 5, which stated that later generation targets tend to be less effective than the first quantified road safety targets ever set in a country or a region.

Finally, hypothesis 6 stated that targets set in countries that have a high incidence of road accidents are more effective than targets set in countries with a low incidence of road accidents. This hypothesis is, at least partly, supported by the results presented in Table 26. Only three cases were classified as high risk (the two targets set in New Zealand, and the target for Hungary), but for these cases an impressive gain of almost 20% in road safety performance was associated with the targets set. There was, however, no clear difference in road safety performance between medium risk and low risk countries. Risk was assessed in terms of the fatality rate per million inhabitants.

On the whole, the results presented in Table 26 exhibit a systematic pattern which broadly supports the hypotheses that were put forward. Four of these hypotheses (number 1, 2, 3, and 4) are clearly supported, the fifth (5) is not supported, and the sixth (6) is partly supported. The differences in road safety performance associated with having a quantified road safety target are, however, surprisingly small. In most cases, the best performance is less than 10% better than the performance of those countries or local governments that did not set quantified road safety targets. It is only targets set by national governments, in particular in high risk countries, that are associated with an improvement in road safety performance of more than 10% during the period the targets were in force.

The weighted mean difference in road safety performance associated with quantified road safety targets amounts to 5.6%, applying to mean period of 7.47 years. This corresponds to a difference of just 0.77% per year, which is certainly well within the range of year-to-year random fluctuations in the number of road accident fatalities in all countries, including those counting more than 40,000 fatalities per year. It is only because the data have been aggregated over both years and countries that this small difference reaches statistical significance. As soon as the data are broken down into subgroups, the difference in road safety performance tends to be statistically significant only in groups for which the difference in road safety performance is greater than the weighted mean difference for the whole data set.

There are, as will be discussed more extensively in Chapter six, a number of problems involved in aggregating data the way it has been done in this analysis. One of these problems is related to the fact that multiple estimates of effect for quantified road safety targets in the same country are not statistically independent. Consider the Netherlands, for example. Two quantified road safety targets set in the Netherlands are represented in the data set. Both were evaluated using Belgium as a comparison. The before and after periods used to evaluate these two targets overlap considerably. It is obvious that the effects attributed to the two targets are statistically dependent, relying as they do, on almost the same data set. This statistical dependency applies to many of the successive targets set in Finland, as well. By combining statistically multiple estimates of effect that are dependent, the statistical precision of the combined mean estimate of effect becomes spuriously high.

In order to reduce this problem as far as possible, the analysis was redone, once again relying on a random effects model, and omitting the following targets:

1. The second generation of local targets set by Norwegian counties (3 observations omitted)

2. The second through fifth quantified road safety target set in Finland, retaining only the first and sixth (4 observations omitted)
3. The first target set in the Netherlands (1 observation omitted)
4. The first target set in New Zealand (1 observation omitted)
5. The second target set in Sweden (1 observation omitted)

In general, the targets for which the shortest periods of data were available were omitted. This reduced the data set from 35 to 25 observations. Table 27 presents the results with the listed observations omitted from the analysis.

Table 27: Weighted mean effects of quantified road safety targets. Random effects model. Ten observations omitted

			Effects attributed to quantified road safety targets	
Hypothesis tested	**Values of variable**	**Number of targets**	**Mean odds ratio**	**95% CI**
1: Target vs no target	Target set	25	0.943	0.894; 0.994
2: Level of ambition	>-5%/year	7	0.870	0.789; 0.960
	-2-5%/year	14	0.954	0.887; 1.026
	<-2%/year	4	1.037	0.915; 1.174
3: Long vs short term	10- years	7	0.847	0.768; 0.935
	1-9 years	18	0.985	0.925; 1.049
4: National vs local	National	15	0.917	0.854; 0.984
	Local	10	0.978	0.902; 1.060
5: First vs revised	First target	17	0.959	0.906; 1.015
	Later targets	3	0.842	0.726; 0.977
6: High vs low risk	High risk	2	0.701	0.572; 0.859
	Medium risk	8	0.969	0.880; 1.068
	Low risk	5	0.924	0.818; 1.043
	Local targets	10	0.978	0.902; 1.060

The results of the analysis of the reduced data set are, for all practical purposes, the same as those based on the complete data set. Hypotheses 1-4 are still supported, hypothesis 5 is rejected, whereas the results with respect to hypothesis 6 are less clear. The magnitude of the difference in road safety performance associated with quantified road safety targets hardly changes when 10 observations are omitted from the analysis. The results of the tests for statistical significance also remain unchanged.

One should note, however, that some categories now contain very few cases, less than five. Results based on less than five cases have an added element of uncertainty, not reflected in the confidence interval, owing to the fact that the extent to which so few cases are representative of a larger, theoretical population is unknown.

5.4 Multivariate analyses of effects of quantified road safety targets

5.4.1 Overview of models fitted

A total of fifteen models were fitted by means of multivariate analysis. These fifteen models differed with respect to the following items:

1. Technique for estimation: Three techniques were applied: (a) Ordinary least squares, assuming normal residuals, (b) Poisson regression, assuming Poisson residuals, and (c) Poisson regression, assuming negative binomial residuals.
2. Correction for heteroskedasticity of residuals: This was applied in some of the ordinary least squares models relying on count data, but not all of them. This correction is part of the mathematical structure of Poisson regression.
3. Set of variables included: Analyses differed with respect to which explanatory variables were included. There were two main classes of independent variables: (a) Country dummies, and (b) Other variables.
4. Set of target variables included: Some analyses included all three target variables that were defined (existence of a target, target level of ambition, and target longevity), other analyses included just one of them.
5. Definition of dependent variable: Three definitions of the dependent variable were used: (a) Fatality count, (b) Fatality rate (population fatality rate), and (c) Logarithm of the number of fatalities.
6. Number of countries included: Some analyses omitted Iceland and Luxembourg, to eliminate instability in fatality rates based on a low number of fatalities.

The following countries that were members of IRTAD as of 1998 were not included, because their time series of data did not go back to 1970:

Czech Republic: Data were available only after 1980.

Germany: Has existed as one country only after 1990.

East Germany: Data before 1990 are incomplete.

Great Britain: Data for the United Kingdom have been used instead.

Hungary: Data were available only after 1980.

Poland: Data were too incomplete for analysis.

South Korea: Data were too incomplete for analysis.

Turkey: Data were too incomplete for analysis.

23 countries with complete data from 1970 through 1998 were retained for analysis, resulting in a data set of 667 observations. West Germany was retained and included in the analyses.

The following countries were coded as having a national quantified road safety target:

Australia, since 1993

Denmark, since 1989

Finland, since 1973

France, since 1997

Iceland, since 1997

Netherlands, since 1986

New Zealand, since 1990

Norway, during the years 1987-1993

Spain, since 1993

Sweden, since 1990

United Kingdom, since 1986

United States, since 1997

A total of 105 observations had targets, 562 did not.

Table 28 gives an overview of the differences between the fifteen models that were fitted.

Models 1, 2 and 3 are identical, except for the set of potentially confounding variables controlled for. Model 4 is identical to model 1, except for correcting for heteroskedasticity in the residuals. It was assumed that such a correction need not be made for the models using fatality rate as the dependent variable, since fatality rate is not a count variable. Models 5, 6 and 7 are identical to models 10, 11, and 12, respectively, except for the omission of Iceland and Luxembourg in models 5, 6, and 7. Models 14 and 15 are identical, except for the assumption made in model 15 that the residuals are distributed according to a negative binomial distribution, whose parameters are estimated as part of the model fitting.

Which of these models, if any, should be preferred? A good model ought to:

1. Include as many explanatory variables as possible, that is all potentially confounding variables and all target variables.
2. Conform to commonly accepted assumptions with respect to the distribution of the residuals.
3. Include all countries.

According to these criteria, models 4, 10 and 15 are the preferred models. However, the criteria for selecting models are not very strict. Hence results from all models will be presented.

Table 28: Overview of multivariate models fitted to estimate the effects of quantified road safety targets. Preferred models in italics

Model	Estimation technique (§)	Correction for hetero-skedasticity	Confounding variables included (#)	Target variables included (*)	Dependent variable	Countries included
1	OLS	No	All	All	Fatalities	23
2	OLS	No	Countries	All	Fatalities	23
3	OLS	No	Others	All	Fatalities	23
4	*OLS*	*Yes*	*All*	*All*	*Fatalities*	*23*
5	OLS	No	All	All	Fatality rate	21
6	OLS	No	All	Target only	Fatality rate	21
7	OLS	No	All	Ambition only	Fatality rate	21
8	OLS	No	Others	All	Fatality rate	21
9	OLS	No	Countries	All	Fatality rate	21
10	*OLS*	*No*	*All*	*All*	*Fatality rate*	*23*
11	OLS	No	All	Target only	Fatality rate	23
12	OLS	No	All	Ambition only	Fatality rate	23
13	OLS	No	Others	All	Fatality rate	23
14	Poisson	Yes (&)	All	All	ln(fatalities)	23
15	*Poisson*	*Yes (&)*	*All*	*All*	*ln(fatalities)*	*23*

(§) OLS = ordinary least squares, Poisson = Poisson regression

(#) Confounding variables include country dummies and other variables (except country dummies)

(*) Target variables include one dummy variable for the presence of a quantified road safety target, one variable representing the level of ambition of a quantified target, and one variable counting the number of years a target has been in force

(&) Model 14 assumes Poisson residuals, model 15 assumes negative binomial residuals

5.4.2 Results by model

Table 29 presents the results according to each model fitted. The results are stated in terms of the odds ratio associated with the target variables, and the lower and upper 95% confidence limits of the odds ratios. The results based on the models that were identified as the preferred ones in section 5.4.1 have been written in boldface italics. All odds ratios have been estimated at the mean value of the dependent variable. The models differ with respect to which confounding variables were included.

The results are difficult to interpret. In most models, the estimated effects of the target variables are not statistically significant. Moreover, the sign of the effects is inconsistent across models. There are 13 estimates of the effect of the target dummy (indicating whether or not a quantified road safety target exists). Eight of these estimates indicate that having a quantified road safety target improves safety performance, five of them indicate the opposite. As far as the ambition of a target is concerned, results refer to an increase of one percentage point in the targeted annual fatality reduction. Estimates of the effect of this variable are generally close to zero and are in the expected direction in only four of thirteen estimates.

Table 29: Effects attributed to target variables in multivariate models. Adjusted odds ratio and 95% confidence interval

		Effect attributed to quantified road safety target		
Model	**Target variable**	**Odds ratio**	**95% CI**	
1 (Fatalities)	Target dummy	1.145	1.021	1.268
	Ambition of target	0.984	1.004	0.963
	Duration of target	0.992	0.977	1.008
2 (Fatalities)	Target dummy	0.884	0.727	1.041
	Ambition of target	0.993	1.019	0.966
	Duration of target	0.992	0.972	1.011
3 (Fatalities)	Target dummy	0.921	0.675	1.168
	Ambition of target	1.030	1.065	0.996
	Duration of target	0.991	0.959	1.023
4 (Fatalities)	Target dummy	***1.145***	***1.066***	***1.223***
	Ambition of target	***0.984***	***1.002***	***0.965***
	Duration of target	***0.992***	***0.984***	***1.001***
5 (Fatality rate)	Target dummy	0.993	0.896	1.090
	Ambition of target	1.001	1.017	0.985
	Duration of target	1.010	0.997	1.022
6 (Fatality rate)	Target dummy	1.036	0.980	1.092
7 (Fatality rate)	Ambition of target	1.005	1.016	0.994
8 (Fatality rate)	Target dummy	0.800	0.636	0.965
	Ambition of target	1.011	1.034	0.988
	Duration of target	0.980	0.959	1.002
9 (Fatality rate)	Target dummy	0.697	0.560	0.834
	Ambition of target	1.002	1.025	0.979
	Duration of target	0.989	0.972	1.006
10 (Fatality rate)	Target dummy	***1.002***	***0.903***	***1.101***
	Ambition of target	***1.001***	***1.018***	***0.985***
	Duration of target	***0.991***	***0.979***	***1.003***
11 (Fatality rate)	Target dummy	1.042	0.986	1.099
12 (Fatality rate)	Ambition of target	1.006	1.017	0.995
13 (Fatality rate)	Target dummy	0.805	0.636	0.975
	Ambition of target	1.010	1.034	0.986
	Duration of target	0.983	0.961	1.005
14 (Ln fatalities)	Target dummy	0.962	0.961	0.964
	Ambition of target	1.005	1.005	1.005
	Duration of target	0.996	0.996	0.997
15 (Ln fatalities)	Target dummy	***0.974***	***0.951***	***0.997***
	Ambition of target	***1.002***	***1.006***	***0.998***
	Duration of target	***0.995***	***0.992***	***0.999***

There are eleven estimates of the effect of the curation variable. This variable describes the number of years since a target was adopted. It takes the value of 1 in the first year, 2 in the second year, and so on. The coefficient for this variable has the expected sign in ten out of eleven cases. The effect attributed to it is in general quite small.

5.4.3 Synthesis of results of multivariate analyses

Is it possible to synthesise the results of these multivariate analyses, with a view to creating a firmer basis for substantive conclusions? As discussed in Chapter 4, a formal synthesis of results based on multivariate models that differ in a number of important respects is not very meaningful. One should rather try to identify which, among the fifteen models tested, that is most credible.

A closer look at the ordinary least squares models, in particular models 1, 2, and 3, which differ only in terms of the variables included, reveals a collinearity problem. In particular, the country dummy variables are almost collinear with other independent variables, especially population and the number of motor vehicles. Inclusion of the United States in the model contributes substantially to this problem. The annual count of fatalities and motor vehicles in the United States is almost an order of magnitude greater than in the second largest country included. Hence, the dummy variable used to identify the United States will be highly correlated with these variables.

The collinearity problem is less severe in the models that do not include the country dummies. Models 3, 8, and 13 do not include the country dummies. The results in these three models are consistent as far the effects of the target variables are concerned. But even these variables are highly correlated.

It is difficult to reach any clear conclusion on the basis of the multivariate analyses. If the results of models 3, 8, and 13 are emphasised, then hypothesis 1 would seem to be supported. Hypothesis 2, concerning the level of ambition of a quantified target, is not supported. Hypotheses 3, on the longevity of a quantified target, is supported. Hypothesis 4, 5 and 6 were not tested in the multivariate analyses.

5.5 Synthesis of results of before-and-after studies and multivariate analyses

The results of the before-and-after studies agree partly with the results of the multivariate analyses. The latter analyses did not add much to the results of the before-and-after studies. Only three of the six hypotheses were tested in the multivariate analyses. The findings of the multivariate analyses can, somewhat generously, be taken to support two of these hypotheses, but not the third. Interpretation is difficult, however, because the three target variables defined in order to test the three hypotheses are highly correlated among themselves.

The before-and-after studies, if taken at face value, support four of the six hypotheses that were tested. The effects attributed to quantified road safety targets are small, in most cases less than a 10% difference in road safety performance. This applies both to the before-and-after studies and to the multivariate analyses.

There are a number of anomalies in the results, which necessitate an extensive discussion of study findings. For the moment, the findings indicate that setting a quantified road safety target may slightly improve road safety performance. This finding is more consistent in the before-and-after studies than in the multivariate analyses. On balance, however, the main tendency of the findings of both analyses support the conclusion that road safety performance is improved by setting a quantified target for improving road safety.

6 Discussion

6.1 Two ways of interpreting study findings

The findings of an evaluation study can almost always be interpreted in two ways:

1. A methodological interpretation, which usually takes the form of a critical examination of a study according to methodological criteria. A study employing a weak design and relying on poor data will often be regarded as inconclusive, based on a methodological interpretation. A methodological interpretation often argues for rejecting the findings of a study.
2. A substantive interpretation, which tries to account for the findings of a study in terms of known causal processes or mechanisms. A substantive interpretation usually argues for taking the findings of a study seriously.

Ideally speaking, one would like to rule out methodological interpretations of a study. In non-experimental accident research, this is not possible. In this chapter, the findings reported in Chapter 5 will be discussed from both a methodological and a substantive point of view.

6.2 History repeats itself until it suddenly no longer does so

The guidelines for observational (that is non-experimental) before-and-after studies developed by Hauer (1997) are intended to strengthen the logical basis for such studies, thus making them more rigorous and less vulnerable to methodological criticism. As an example, making the assumption that a comparison group accurately predicts what would have happened in the treated group in the absence of treatment seems more defensible if it can be justified by reference to past predictive success, than if the predictive performance of the comparison group is unknown.

Despite its intuitive appeal, this argument is not logically watertight. It relies on an assumption that no amount of predictive performance in the past can substantiate, namely that history will go on repeating itself in the after period the same way it did in the before period.

Consider the very instructive case presented in Figure 28. It is based on a paper by Susan Partyka (1991). She fitted a very simple model to fatality count data for the United States from 1960 through 1982. The model included just three numerical variables and a dummy variable for the year 1974. As shown by Figure 28, the fitted values trace the actual values as a shadow; in fact the two series (actual and predicted) are almost indistinguishable. The squared correlation coefficient for the model fitted to 1960-1982 data was 0.98.

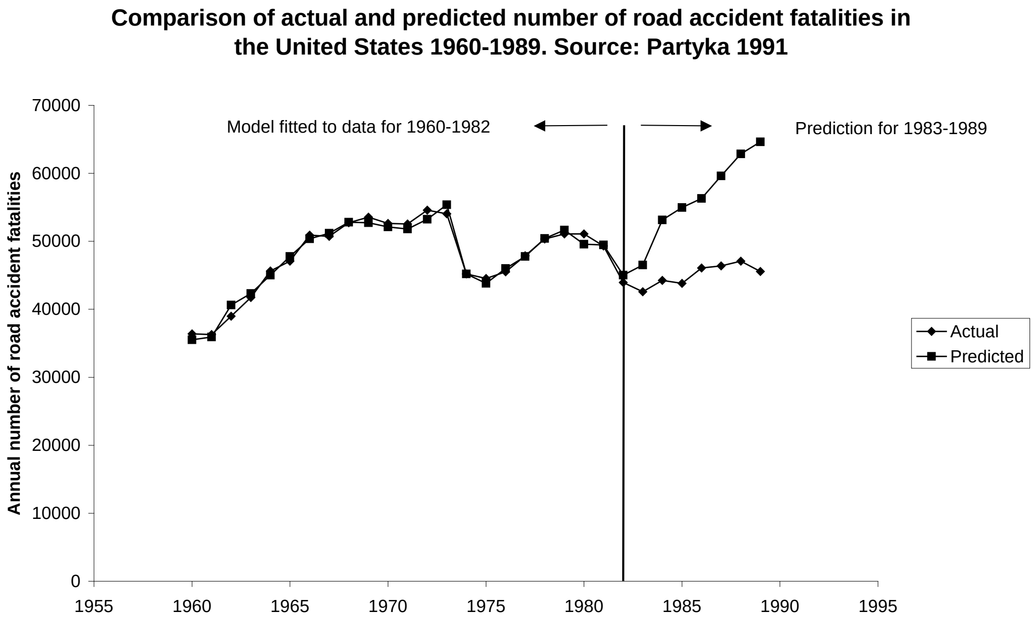

Figure 28: Predictive performance of a model fitted to fatality data for the United States. Based on Partyka 1991.

But look what happened when this model was used to predict for the years 1983-1989. By 1989, it over predicted the number of fatalities by nearly 20,000. The fact that this model fitted past data as perfectly as any model possibly could, did not ensure that it predicted correctly. History repeats itself until it no longer goes on repeating itself. Past predictive success in no way ensures future predictive success.

The simplest assumption one can make in a before-and-after study is that: “Next year will be like this year”. Finding cases in which this simple assumption is wrong is easy. Yet, even going to great lengths in demonstrating past predictive success of a comparison group does not establish the validity of assuming that future predictions will be as successful as past. One always relies on a non-testable assumption that the future will be like the past.

6.3 Alternative analyses of before-and-after studies

The before-and-after studies of quantified road safety targets set by national governments, reported in Chapter 5, did not utilise all information contained in the data. Fatality counts were summed for all years before and all years after, yielding a 2 x 2 table based on the totals in the before period and after period for the target country and the comparison country. By adding data for several years this way, long term trends are masked, and the analysis will not control for them. On the other hand, adding data for several years makes the estimate more precise, by increasing the size of the accident sample. Moreover, an estimate was sought for the effect of a quantified road safety target referring to the entire period the target had been effective, not a single year of that period.

However, as noted in the presentation of the results for Norway in Chapter 5, the results of the 2 x 2 table analyses are sometimes anomalous. In the case of Norway, all observations were consistent with the hypothesis that the target was effective in improving safety performance. Despite this, the estimate of effect based on the 2 x 2 table indicated a net increase of the number of road accident fatalities after the target was set.

Alternative ways of analysing the data collected for the before-and-after studies have therefore been explored. One possible approach to analysis is to project the long-term trend observed in the before-period in the target country to the after period. This approach does not utilise information about the comparison country. The long-term trend in the before period was described in terms of the mean annual percentage change of the number of fatalities.

Consider Australia as an illustration. There are eight years of before data (1985-1992); annual percentage changes from the previous year are available for seven of these years (all except the first year of the before period). Six of the annual changes were negative, that is the number of fatalities was reduced from the previous year. The mean annual percentage change was –5.3%. The following very simple model of long-term trends was constructed.

The first year of the before period was set equal to the recorded number, that is 2,941 road accident fatalities in Australia in 1985. Each successive year was then reduced by 5.3%. Thus for 1986, the estimate was 2,941 x 0.947 = 2,785. For 1987, it was 2,785 x 0.947 = 2,638, and so on. The trend was projected to the after period. The actual number of fatalities in the after period was compared to the projected number (corrected for inaccuracy of prediction, as observed in the before period). The effect of the target was estimated in terms of the ratio of the actual to the predicted number of fatalities.

This approach to analysis takes care of long term trends, but not of year-to-year fluctuation in the number of road accident fatalities.

Another possible approach to analysis is to carry the odds ratio method employed in the before period forward to the after period. The observed number of fatalities in the target country in the last year of the before period is taken as basis. A prediction for the after period is obtained by relying on the year to year changes in the *comparison* country in the after period.

Once more using Australia as an illustration, the fatality count in the last year of the before period (1992) was 1,974. In the first year of the after period, the number of fatalities in West Germany, used as the comparison country, declined by 5.1%. If the same decline had occurred in Australia, the predicted number of fatalities for 1993 would have been 1,974 x 0.949 = 1,873. Similar predictions were made for each year of the after period, always using the last year of the before period in the target country as basis.

These two approaches to analysis have been applied to all countries that were included in Chapter 5. Table 30 shows the results.

Table 30: Results of alternative approaches to analysis in before-and-after studies

		Estimates of effect of target based on alternative approaches to analysis – odds ratio		
Country	**Target**	**Projected trend in target country**	**Annual changes in comparison country**	**Matched odds ratio in before period**
Australia	1	1.094	0.988	0.971
Denmark	1	0.884	0.819	1.165
Finland	1	0.679	0.732	0.997
	2	0.998	0.826	0.805
	3	0.987	1.001	0.944
	4	1.472	1.088	0.936
	5	1.071	1.031	1.340
	6	0.645	0.765	0.813
France	1	1.073	0.993	1.018
Hungary	1	0.594	0.702	0.589
Iceland	1	1.159	2.240	1.828
Netherlands	1	1.384	0.917	0.859
	2	1.020	0.858	0.890
New Zealand	1	0.862	0.890	0.973
	2	0.603	0.800	0.815
Norway	1	0.794	0.799	1.062
Poland	1	1.122	1.043	1.104
Spain	1	0.572	0.777	0.862
Sweden	1	0.767	0.712	0.752
	2	1.098	0.974	0.973
United Kingdom	1	1.016	0.851	0.950
United States	1	0.950	N.A.	1.020
Mean (simple)	All	0.948	0.943	0.985

N. A. = not applicable

Table 30 includes 22 quantified road safety targets. For each target, three estimates of effect are presented. The one labelled “matched odds ratio in before period” was used in the main analysis, presented in Chapter 5. The other two estimators have been discussed above.

The three estimates agree as far as the direction of impact is concerned in 12 of the 22 cases included. They disagree about the direction of impact in 10 of the 22 cases. The mean of the estimates (simple mean; no statistical weighting) all point in the same direction, indicating that a very small improvement in safety performance is associated with the adoption of a quantified road safety target.

If one takes the mean of the three estimates, including all 22 cases, it comes to 0.959, or a 4% gain in safety performance associated with quantified targets. As argued in Chapter 5, the results for multiple targets set in same country are not independent. If targets 2 through 5 in Finland, target 1 in the Netherlands, target 1 in New Zealand, and target 2 in Sweden are omitted, the mean estimate of effect for all three approaches to analysis comes to 0.936, or about 6% gain in safety performance. A 95% confidence interval for this (unweighted) estimate is from 0.804 to 1.068 (plus or minus 1.96 x standard error of the mean).

6.4 Methodological issues in multivariate analyses

The major methodological issues in the multivariate analyses reported here concern:

1. The potential for omitted variable bias
2. Collinearity among the variables included
3. The functional form of the relationships between variables

There is at least one potentially very important source of systematic error in the multivariate analyses made in this report. It has to do with why a country chooses to adopt a quantified road safety target. The sample used in the multivariate analyses consisted of 667 observations, of which 105 had a quantified target. These 105 data points are, however, by no means a random sample of all 667 data points. The decision to set a quantified road safety target can be influenced by variables included in the model in at least three ways:

1. A quantified target may be set when a country has had a particularly bad year. Regression-to-the-mean would then bring the count of fatalities back to a more normal level, even if the target is ineffective.
2. Quantified targets tend perhaps to be set by countries that are particularly concerned about road safety. These countries might achieve a better safety record than other countries even if no target was set.
3. Quantified road safety targets tend perhaps to be set by countries that feel that they lag behind other countries in improving road safety. These countries may have a greater potential for improving safety than other countries.

All these cases are examples of selective recruitment. The selective recruitment of countries setting quantified road safety targets is a problem for multivariate analysis, because the process of selective recruitment is both influenced by variables in the model, and influencing these variables. However, no way has been found of modelling the process of selective recruitment for the purpose of accounting for it in multivariate analyses.

What this means is basically that the direction of causality between targets and safety performance cannot be determined in the multivariate analyses. It could be the case that – due to selective recruitment – there is an association between safety performance and target setting, such that, for example, countries with a comparatively good safety record set demanding quantified road safety targets. It could also be the other way around: that setting targets causes safety performance to be improve.

Collinearity is a severe problem. Figure 29 gives an example of this. It shows the relationship between the count of road accident fatalities and the dummy variable identifying the United States of America.

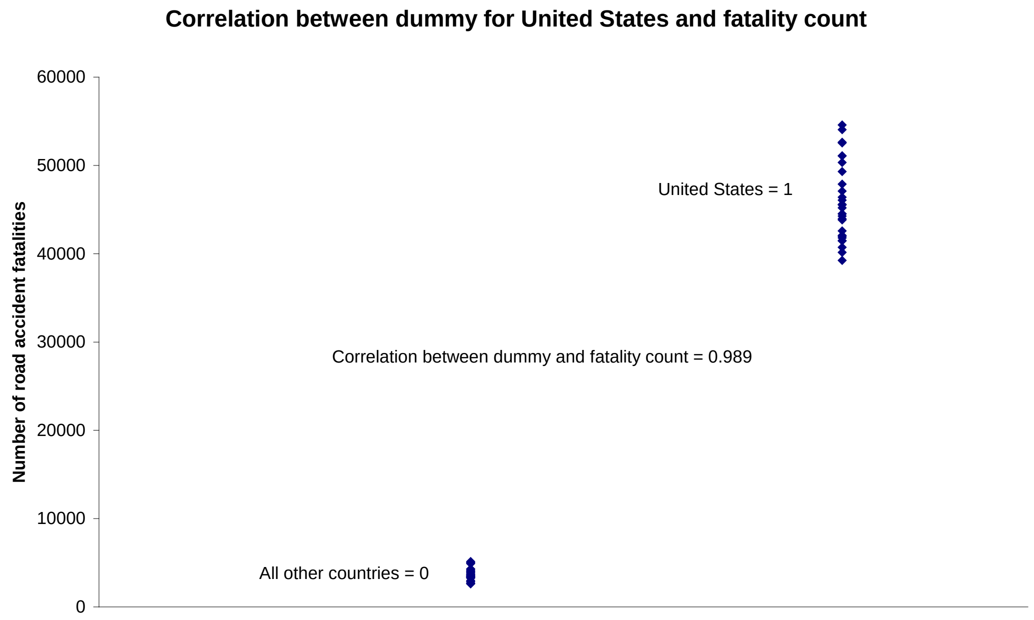

Figure 29: Correlation between country dummy for the United States and fatality count

The correlation between these two variables is 0.989, that is virtually 1. It is therefore nearly impossible to estimate the effects on fatality counts of any other variable very precisely in analyses that include the United States.

On the other hand, it was regarded as desirable to include as many countries as possible in the analysis. The United States has recently adopted a quantified road safety target. Being by far the biggest motorised country, the experience of the United States in relying on a quantified road safety target is of particular interest. In general, as many countries as possible were included in the analyses, in the belief that countries can learn from each other as far as road safety policy is concerned.

Analyses were made in which the two smallest countries in the data set, Iceland and Luxembourg, were excluded, since their annual fatality counts are so low as to be very much influenced by random variation. The results, however, were still very uncertain.

Finally, a few words on the issue of non-linear effects. In most of the analyses, ordinary least squares linear regression has been applied. A logarithmic specification of the relationship between variables was applied in some models. In general, the linear specification is adequate. Figure 30 provides an illustration of this. It shows the relationship between population and fatality counts in the 23 countries included in the multivariate analyses.

The linear model fits these data very well (R-squared = .9021). In this case a polynomial fitted the data slightly better (R-squared = .9099), as did a geometric function (R-squared = .9398). The fit of the linear model is, however, adequate by any reasonable standard. The residuals do, however, tend to fan out in the diagram, as indicated by the dotted lines. This phenomenon is known as heteroskedasticity, and it was adjusted for in the models that used fatality counts as the dependent variable.

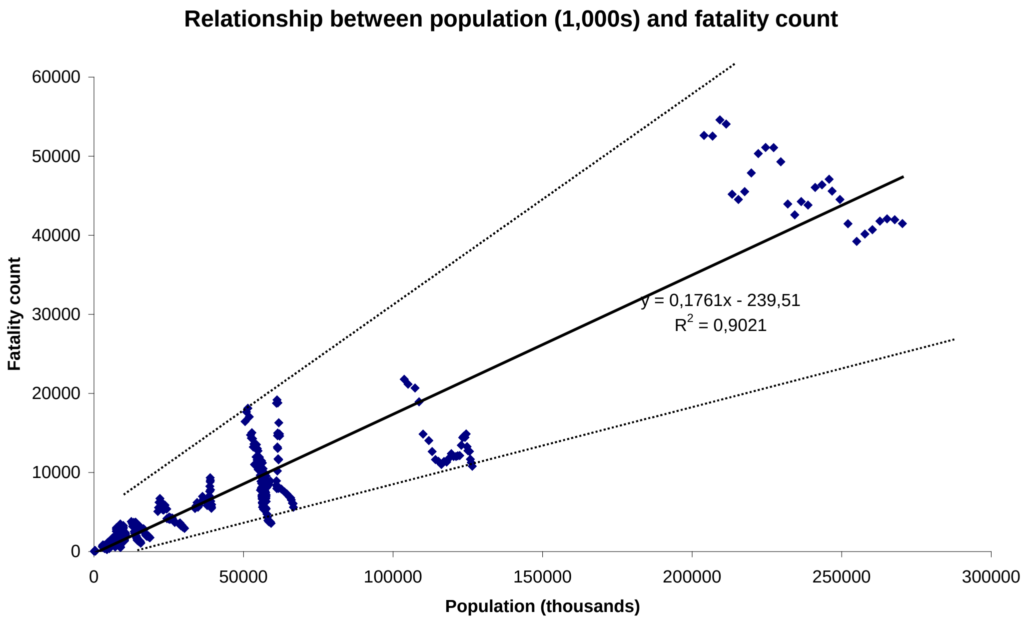

Figure 30: Fit of a linear model of the relationship of population to fatality counts

The most severe problems for the multivariate analyses are selective recruitment of countries setting quantified road safety targets and collinearity among the explanatory variables. These problems are so severe as to render the results of the multivariate analyses very difficult to interpret.

In short, the multivariate analyses, although superior to before-and-after studies in controlling for more confounding factors, do not seem to add much of value to this study. These analyses are, regrettably, likely to be influenced by both omitted variable bias and the problem of collinear explanatory variables.

6.5 Internal validity of study findings (causal interpretation)

Next to statistical validity, perhaps the most important aspect of study validity in road safety evaluation research is internal validity. Internal validity denotes the extent to which one can infer a causal relationship between the safety programme of interest and changes in safety performance. In short: Were the observed changes in road safety performance caused primarily by the setting of quantified road safety targets, or were they caused primarily by other factors?

To answer this question, the internal validity of the study presented in Chapter 5 has been assessed in terms of nine criteria for causality in an observed statistical relationship. These nine criteria are:

1. There should be a statistical relationship between the presumed cause and the presumed effect.
2. A strong statistical relationship is, keeping everything else constant, more likely to be causal than a weak statistical relationship.
3. The statistical relationship should be internally consistent in subsets of the data.

4. The direction of causality should be clear, that is it should be clear which variable is the cause and which is the effect.
5. The statistical relationship between cause and effect should not disappear when confounding factors are controlled.
6. If the cause comes in different doses, there should be a dose-response pattern between cause and effect.
7. If the cause can reasonable be assumed to be effective only within a certain subset of the data, effects should be found only in that subset and not outside of it (the specificity of effect criterion, to be elaborated below).
8. The causal mechanism through which effects are transmitted should be known.
9. The findings of the study should be supported by theory or evidence from other, preferably rigorous, studies.

These points will be discussed in turn.

6.5.1 Existence of statistical relationship

The summary estimate of effect in the before-and-after study, based both on the complete data set (35 cases) and on the data set in which 10 targets were omitted in order to ensure the statistical independence of the individual estimates of effect, was statistically significant at the 5% level. A statistical relationship therefore exists between the causal variable (targets) and the outcome variable (safety performance) in the before-and-after studies.

As far as the multivariate analyses are concerned, thirteen estimates were made of the target dummy variable. Five of these found a statistically significant favourable impact on road safety performance, three found a statistically non-significant favourable impact on road safety performance. The other five estimates indicated an adverse impact on road safety performance of setting quantified road safety targets. Two of these five estimates were statistically significant at the 5% level.

It is concluded that a statistical relationship between targets and safety performance has been found, in particular in the before-and-after studies.

6.5.2 Strength of statistical relationship

Ceteris paribus, a strong statistical relationship is generally taken as a stronger indication of causality than a weak statistical relationship. The overall statistical relationship found in the before-and-after studies is weak. The effect attributed to quantified road safety targets is generally less than a 10% difference in safety performance.

In the multivariate analyses, the effect attributed to the target dummy varies between a 15% increase in the number of fatalities and a 30% reduction. Most estimates are close to zero.

It is concluded that the statistical relationship between targets and road safety performance is weak.

6.5.3 Internal consistency of statistical relationship

Internal consistency denotes the extent to which the individual estimates of effect are consistent – in direction and magnitude – with the overall estimate of effect. If, for example, the overall estimate indicates an adverse effect of quantified road safety targets, while the great majority of the individual estimates of effect indicate a favourable effect, internal consistency is low.

A total of 66 estimates of effect were made in the before-and-after study. These 66 estimates can be divided into three times 22 estimates, each set of 22 estimates being based on a different model of analysis. All estimates are presented in Table 30. 10 of the 22 estimates are inconsistent with the overall estimate of effect for the analysis based on projected trend in the target country. 5 estimates of effect are inconsistent with the overall estimate of effect for the analysis based on year-to-year changes in the comparison country in the after period. 7 estimates of effect are inconsistent with the overall estimate in the main analysis, based on the matched odds ratio method, presented in Chapter 5. The percentage of individual estimates of effect that are consistent with the mean estimate, taking all 66 estimates together, is 67%.

Thirteen estimates of the effect of a quantified road safety target were made in the multivariate analyses. Eight of these indicated a gain in safety performance associated with a target, five indicated a loss. 62% of the estimates are consistent with the hypothesis that setting a quantified road safety target improves safety performance.

It is concluded that the majority of individual estimates of effect are consistent with the summary estimate of effect. The percentage of inconsistent findings is, however, disturbingly large. The level of consistency in study findings is below the level of reliability generally regarded as acceptable in repeated measurements of the effects of a given variable.

6.5.4 Clarity with respect to direction of causality

In most non-experimental research, causal relationships are generally assumed to go in one direction only, from cause to effect, and not in both directions at the same time, or in the opposite direction, from effect to cause.

As noted in discussing the results of the multivariate analyses, it is not difficult to imagine how causality could be reversed in this case. One can imagine a number of ways in which the count or rate of road accident fatalities – the dependent variable in this study – could influence the decision to set a quantified road safety target – the primary causal variable of interest in this study. Three possibilities were mentioned above.

The first of these, an abnormally high count of fatalities in the period immediately before a target was set, will be discussed below, as a potential confounding factor in the study. The other two ways in which prior accident history could influence target setting, were by the selective recruitment of either the safest countries or the least safe countries to those who set targets. Neither possibility can be ruled out.

Table 31 compares the public health risk attributable to road accident fatalities during the before period in the target and comparison countries. Risk is stated in terms of the number of road accident fatalities per million inhabitants in the before period (mean for all years combined).

Table 31: Health risk attributable to road accident fatalities in the before period. Target and comparison countries

Target no	Target country	Road accident fatalities per million inhabitants	Comparison country	Road accident fatalities per million inhabitants	Possible selection bias?
1	Australia	156	West Germany	129	Yes
1	Denmark	141	Switzerland	169	Yes
1	Finland	235	Denmark	229	No
2	Finland	195	Denmark	178	No
3	Finland	201	Denmark	203	No
4	Finland	180	Denmark	179	No
5	Finland	120	Switzerland	173	Yes
6	Finland	123	Japan	105	No
1	France	156	Italy	122	Yes
1	Hungary	184	Czech republic	111	Yes
1	Iceland	59	Luxembourg	175	Yes
1	Netherlands	152	Belgium	238	Yes
2	Netherlands	139	Belgium	226	Yes
1	New Zealand	228	Canada	159	No
2	New Zealand	216	Japan	102	No
1	Norway	97	Switzerland	171	Yes
1	Poland	181	Portugal	233	Yes
1	Spain	203	Portugal	290	Yes
1	Sweden	98	United States	198	Yes
2	Sweden	76	Canada	126	Yes
1	United Kingdom	116	France	259	Yes
1	United States	157	None found	N.A.	Unknown

The choice of comparison country for each target country was made in this study. It is not known whether the countries that have set targets compared themselves and their road safety record to that of other countries in the process of setting targets. The data presented in Table 31 are therefore an imperfect indication of the possible presence of selective recruitment bias.

A possible selective recruitment bias is indicated whenever the population fatality rates from road accidents differ between the target and comparison country by more than 10% in either direction. As shown in Table 31, the selective recruitment of a particularly safe or particularly unsafe country to those who choose to adopt a quantified target cannot be ruled out in 14 of the 22 cases included.

Although the limitations inherent in these comparisons must not be forgotten, they at the very least show that the possibility of a reverse direction of causality cannot be ruled out. This conclusion applies *a fortiori* to the multivariate analyses.

6.5.5 Control of confounding factors

One of the most important characteristics of an observational road safety evaluation study, is the degree to which it controls for potential confounding variables. In theory, a multivariate analysis is better than a conventional before-and-after study in this respect. In the present study, however, the results of the multivariate analyses appear to be strongly influenced by collinear variables. Coefficients estimated for the target variables vary erratically between model specifications and do not provide results that are easily interpretable.

Turning to the before-and-after studies, one potential confounding variable which has not been explicitly controlled for, is regression-to-the-mean. This confounding variable could influence study results if the count of fatalities in the year or years immediately before a quantified target was set was abnormally high or low compared to the long term trend. Long term trends, and implicitly any factor affecting them, were controlled for in the analyses in which trends from the before period were projected to the after period.

The possible presence of regression-to-the-mean has been assessed by comparing the actual number of road accident fatalities in the last year of the before period to the number predicted on the basis of long term trends for the 22 quantified road safety targets set by national governments. If these two numbers differed by more than 10%, it was judged that the possibility of regression-to-the-mean could not be ruled out. The only case in which regression-to-the-mean could not be ruled out, was Iceland. The recorded number of fatalities for Iceland in the last year before a target was set, was abnormally low. Hence, part of the increase in the number of fatalities that was found during the after period may have been regression-to-the-mean.

Apart from this case, regression-to-the-mean is unlikely to have affected the results of the before-and-after studies. It is concluded that, although control of confounding factors was imperfect in both the before-and-after studies and the multivariate analyses, it is difficult to identify a specific potentially confounding factor that ought to have been controlled, but was not. This conclusion rests on the assumption that the effects of, for example, turns of the business cycle, are captured in the long term trends in fatality counts, since fluctuations of the business cycle are one of the foremost factors influencing long term trends in road safety. The multivariate analyses explicitly controlled for annual real growth of national income.

6.5.6 Presence of a dose-response pattern

The main causal variable of interest in this study, quantified road safety targets, comes in different doses in at least two respects:

1. Some targets are more ambitious than others, aiming for a larger reduction of the number of road accident fatalities.
2. Some targets are more long term than others, giving more time to introduce effective road safety measures.

According to the before-and-after studies, both these characteristics of quantified road safety targets are associated with a dose-response pattern in the expected direction. The results of the multivariate analyses are less clear. The coefficients for the variable intended to measure the ambitiousness of a target are generally non-significant. The coefficients for the variable intended to measure the length of the period to which a target applies are generally in the expected direction, but not statistically significant at conventional levels.

The three target variables included in the multivariate analyses are, however, highly correlated among themselves, making it difficult to estimate the partial effect of each of them very precisely. Correlations range from –.63 to –.83.

It is concluded that the before-and-after studies provide some evidence of a dose-response pattern. Evidence from the multivariate analyses is inconclusive.

6.5.7 Specificity of effect in subset of data

This criterion is relevant if, for example, targets have been set specifically for certain groups of road users, or certain locations, but not for others. This is not the case for any of the targets included in this study. A case in point would be to set a more demanding target for reducing accidents involving young drivers, for example, than the target set for reducing the total number of accidents or accident victims.

No case of such sub-targets, amenable to evaluation, have been found in this data set. Hence this criterion of causality is not applicable.

6.5.8 Knowledge of causal mechanisms

A quantified road safety target does, by itself, not directly influence road safety. Such a target can only be effective if it leads to the adoption of an effective road safety programme. This causal mechanism is indicated below:

Adoption of a quantified road safety target	➔	Implementation of an effective road safety programme	➔	Reduction of the number and severity of road accidents

Unfortunately, this causal mechanism cannot be described very well, if at all, for most of the quantified road safety targets that have been evaluated in this report. An attempt was made to describe this causal chain in an earlier evaluation of quantified road safety targets set by Norwegian counties (Elvik 1993). It was found that counties that adopted very ambitious targets did increase their spending on road safety programmes more than other counties, at least during the first of the two long term planning periods that were considered. During the last of these two periods, the relationship between targets and spending on road safety programmes was less clear cut.

The present study would have been strengthened if it had been found that the countries setting very demanding targets were also more successful than other countries in introducing effective road safety programmes. Showing this, would, however, require fairly detailed data concerning road safety programmes that were actually implemented.

Recent evaluations of road safety policies in Norway (Elvik 1999) and Sweden (Elvik and Amundsen 2000) have found that current policy priorities are rather inefficient in both countries (Elvik 2001). If this applies to other countries as well, it may go some way towards explaining why even targets aiming for a drastic reduction of road accident fatalities do not appear to be very effective.

It is concluded that evidence of the causal mechanism by which road safety targets can influence road safety has not been presented in this study.

6.5.9 Results reproduced in similar studies

Confidence in the results of an evaluation study is enhanced if other studies confirm the results, or if the results can be shown to be consistent with well established theory. Neither of these sources of confirmation can lend very much support to the results of this study. As noted in Chapter 2, there are few previous studies. The findings of these studies are not perfectly consistent, and some of the previous studies were not very rigorous from a methodological point of view. At best, these studies provide weak evidence on the effects of quantified road safety targets.

Six hypotheses were proposed. These hypotheses do not constitute a strong body of theory, but are nevertheless useful as a reference in interpreting study findings. Four of the six hypotheses were supported by the before-and-after study, one was rejected, and evidence was inconclusive with respect to the sixth. Only three of the hypotheses were tested in the multivariate analyses. Two were supported, one was not.

6.5.10 Conclusions with respect to internal validity

Table 32 summarises the results of the discussion of the internal validity of this evaluation study.

The impression gained from the overview in Table 32 is mixed. Some of the criteria of internal validity are fairly well supported in this study, other criteria are not at all supported, or not applicable. Based on this assessment, it must be concluded that a methodological interpretation of study findings cannot be ruled out. It has not been possible to implement a sufficiently rigorous study to rule out methodological artefacts.

On the other hand, there does seem to be a certain systematic pattern in the results, at least of the before-and-after study. If the results of this study are taken at face value, they do indicate that setting a quantified road safety target will, on the average improve road safety performance slightly. The effect, if real, is however very small, less than 1% net reduction of the number of road accident fatalities per year. In most countries, this reduction is smaller than the annual reduction many of these countries experienced long before any quantified road safety target was set.

Table 32: Summary of discussion of internal validity

Criterion	Assessment
1: Existence of statistical relationship	Yes, in before-and-after studies. Doubtful in multivariate analyses
2: Strength of statistical relationship	Weak in both before-and-after study and multivariate analyses
3: Consistency of statistical relationship	67% of individual findings consistent with summary estimate in before-and-after study
4: Direction of causality	Reverse causality cannot be ruled out – very difficult to design a study that would rule out this possibility
5: Control of confounding factors	Incomplete in both before-and-after study and multivariate analyses
6: Presence of dose-response pattern	Present in before-and-after study; doubtful in multivariate analyses
7: Specificity of effect in sub group	This criterion is not relevant for this study
8: Description of causal mechanism	No adequate description of the causal mechanism of quantified road safety targets could be given
9: Findings reproduced in other studies	There are few other studies. The majority of the hypotheses proposed were supported

7 Conclusions

The idea of setting quantified targets for improving road safety has gained increasing support in recent years, as evidenced by the fact that more and more countries have set such targets. It is therefore of some importance, and obviously of high relevance for policy makers, to find out whether setting quantified road safety targets leads to more effective road safety programmes, which in turn leads to better safety performance.

In this report, an attempt has been made to evaluate the effects on road safety performance of quantified targets set both by national governments in many countries and by local governments in a few countries. The effects of these targets were evaluated by means of a set of before-and-after studies, and by means of multivariate analyses, employing data from 23 countries covering a period of 29 years (for a total sample size of 23 x 29 = 667).

Neither the before-and-after studies nor the multivariate analyses were very successful in providing clear evidence of the effectiveness of quantified road safety targets. It turned out to be very difficult to implement these study designs in a sufficiently rigorous manner to rule out methodological explanations of study findings. Hence the main conclusion of this study is:

The evaluation of the effects on road safety performance of quantified road safety targets is inconclusive. It is not possible to place confidence in study findings, as these could reflect anomalies in the data, effects of variables not controlled for, a reversed direction of causality, or inadequate techniques for statistical analysis, rather than the true effects of quantified road safety targets.

If, in spite of this, one decides to take study findings at face value, the before-and-after studies show that quantified road safety targets are associated with a small gain in road safety performance – on the average amounting to less than 1% per year. The findings of the multivariate analyses are too messy to justify any substantive conclusion regarding the effects of quantified road safety targets.

8 References

Breen, J. Visions, targets and strategies. Paper presented at European Road Safety Conference, 16-18 April 2000, Amsterdam. Brussels, ETSC, 2000.

Commission of the European Communities. Priorities in EU road safety. Progress report and ranking of actions. COM (2000) 125 final. Brussels, March 17, 2000.

Elvik, R. Quantified road safety targets: a useful tool for policy making? Accident Analysis and prevention, 25, 569-583, 1993.

Elvik, R. Evaluations of road accident black spot treatment: a case of the Iron Law of evaluation studies? Accident Analysis and Prevention, 29, 191-199, 1997.

Elvik, R. Bedre trafikksikkerhet i Norge. TØI rapport 446. Oslo, Transportøkonomisk institutt, 1999.

Elvik, R. Setting goals – effective road safety target setting. Proceedings of Best in Europe Road Safety Conference, Brussels 12 September 2000, 10-18. Brussels, European Transport Safety Council, 2000A.

Elvik, R. Quantified road safety targets : an international assessment of effectiveness. Proceedings of IRTAD Seminar in Vienna, Austria, September 25-26, 2000B (Available on CD-rom).

Elvik, R. Improving road safety in Norway and Sweden: analysing the efficiency of policy priorities. Traffic Engineering and Control, January 2001, 9-16.

Elvik, R. The importance of confounding in observational before-and-after studies of road safety measures. Accident Analysis and Prevention, forthcoming, 2002.

Elvik, R.; Amundsen, A. H. Improving road safety in Sweden. Main report. Report 490. Oslo, Institute of Transport Economics, 2000.

Elvik, R.; Mysen, A. B. Incomplete accident reporting: Meta-analysis of studies made in thirteen countries. Transportation Research Record, 1665, 133-140, 1999.

Elvik, R.; Mysen, A. B.; Vaa, T. Trafikksikkerhetshåndbok. Tredje utgave. Oslo, Transportøkonomisk institutt, 1997.

European Transport Safety Council (ETSC). A strategic road safety plan for the European Union. Brussels, ETSC, 1997.

European Transport Safety Council (ETSC). ETSC update: Visions, targets and strategies. Newsletter on road safety management. Brussels, ETSC, May 2001.

Fleiss, J. L. Statistical methods for rates and proportions. Second edition. New York, NY, John Wiley & Sons, 1981.

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

Færdselssikkedskommissionen. Færdselssikkerhedspolitisk handlingsplan. Betænkning nr 1157. København, Statens Informationstjeneste, 1988.

Færdselssikkerhedskommissionen. Færdselssikkerhedskommissionens strategiplan 1995-2000. København, Trafikministeriet, 1996.

Færdselssikkerhedskommissionen. Hver ulykke er én for meget. Trafiksikkerhed starter med dig. Mod nye mål 2001-2012. København, Trafikministeriet, 2000.

Gaudry, M. Multiple levels, damages, forms, moments and variables in road accident models. In Gaudry, M.; Lassarre, S. (Eds). Structural Road Accident Models. The International DRAG Family, 1-36. Oxford, Pergamon Press (Elsevier Science), 2000.

Gaudry, M.; Lassarre, S. (Eds). Structural Road Accident Models. The International DRAG Family. Oxford, Pergamon Press (Elsevier Science), 2000.

Hauer, E. Comparison groups in road safety studies: an analysis. Accident Analysis and Prevention, 23, 609-622, 1991.

Hauer, E. Observational before-after studies in road safety. estimating the effect of highway and traffic engineering measures on road safety. Oxford, Pergamon Press (Elsevier Science), 1997.

Hauer, E.; Ng, J. C. N., Papaioannou, P. Prediction in road safety studies: an empirical inquiry. Accident Analysis and Prevention, 23, 595-607, 1991.

Hungarian Government. National Traffic Safety Programme of Hungary. Budapest, Hungarian government, March 1993.

Jaeger, L.; Lassarre, S. The TAG-1 model for France. In Gaudry, M.; Lassarre, S. (Eds). Structural Road Accident Models. The International DRAG Family, 157-184. Oxford, Pergamon Press (Elsevier Science), 2000.

Land Transport Safety Authority. National Road Safety Plan 1995. Wellington, NZ, Land Transport Safety Authority, 1995

Lebrun, D. The target –10% programme. PTRC Summer Annual Meeting (proceedings of seminar H), 163-169, 1989.

Ministry of Transport and Communications Finland. Road safety plan 1997-2000. The Recommendation of the Consultative Committee on Road Safety. Publications 28/97. Helsinki, Ministry of Transport and Communications, 1997.

National Road Safety Committee. Road safety strategy 2010. A consultation document. Wellington, NZ, Land Transport Safety Authority, 2000.

OECD Scientific Expert Group. Targeted Road Safety Programmes. Paris, OECD, 1994.

OECD Scientific Expert Group. Draft final report on safety management and implementation strategies. Paris, OECD, October 24, 2000.

Partyka, S. Simple models of fatality trends revisited seven years later. Accident Analysis and Prevention, 23, 423-430, 1991.

Risser, R.; Michalik, C. Aktion „Minus 10 Prozent“ in Österreich. Bergisch Gladbach, Forshnungsbericthe der Bundesanstalt für Strassenwesen, Bereich Unfallforshung, Heft 159, 1987.

Schlabbach, K. Erhöhung der Verkehrssicherheit nach Plan. Realistischer Ansatz oder perspektivlose Träumerei? Zeitschrift für Verkehrssicherheit, 36, 146-155, 1990.

Shadish, W. R.; Haddock, C. K. Combining estimates of effect size. In Cooper, H.; Hedges, L. V. (Eds): The Handbook of Research Synthesis, Chapter 18, 261-281. New York, NY, Russell Sage Foundation, 1994.

Statens institut för komunikationsanalys (SIKA). Vidareutveckling av de transportpolitiska målen. SIKA rapport 2000:1. Stockholm, Statens institut för komunikationsanalys, 2000.

US Department of Transportation. The changing face of transportation. Report BTS 00-007. Washington DC, US Department of Transportation, Bureau of Transportation statistics, 2000.

Technical University of Gdansk (corporate author). GAMBIT. Integrated Programme of Road Safety Improvement in Poland. Gdansk, Technical University of Gdansk, 1995.

Traffic Safety Committee. Traffic safety plan 1997-2001. report presented to the Althing by the Minister of Justice on the implementation of the Traffic Safety Plan. Reykjavik, Ministry of justice, 1997.

Transport Canada. Road safety vision 2001. 2000 update. Ottawa, Minister of Public Works and Government Services, 2001.