Never Stand Still | Science | Transport and Road Safety (TARS) Research

An evaluation of the methods used to assess the effectiveness of mandatory bicycle helmet legislation in New Zealand

Joanna Wang
Research Fellow at Transport & Road Safety (TARS)
University of New South Wales

Australasian College of Road Safety Conference
November, 2013

Co-authors

- Prof. Raphael Grzebieta, Transport and Road Safety, University of New South Wales
- Mr. Scott Walter, Centre for Health Systems and Safety Research, University of New South Wales
- Dr. Jake Olivier, School of Mathematics and Statistics, University of New South Wales

Outline

- Mandatory Helmet Legislation in NZ
- Povey et al. (1999) and Robinson (2001)
- Other studies
- Conclusions and recommendations

MHL in NZ

Mandatory Helmet Law (MHL) in NZ

- Came into effect on 1 Jan 1994
- Applies to all age groups for on-road cycling
- Voluntary helmet use had been promoted in NZ prior to law
- Helmet wearing rate:
  - ~0 in 1986
  - 84% (5-12 yrs old), 62% (13-18 yrs old), 39% (>18 yrs old) in 1992
  - > 90% for all age groups after the law
- Solid evidence for helmet wearing in lowering bicycle related head injuries from biomechanical and epidemiological studies
  - Aim of the legislation: increase the helmet wearing rate, in an effort to reduce head injuries to cyclists

Our study

Aim of our study: review and critically evaluate studies assessing the effectiveness of the NZ bicycle helmet law, NOT to assess the law itself

- Focus only on studies that analyse bicycle helmet use and cycling head injuries in NZ context; Scopus and Google Scholar search

Povey et al. (1999)

Povey et al. (1999)

- Hypothesis: 'This paper considers the effect of cycle helmet wearing on hospitalised head injuries.'
- Data are aggregated by year (1990-1996)
- Cyclist limb fractures used as a measure of cycling exposure
- Motor and non-motor vehicle crashes analysed separately
- Non-MVC broken down into three age groups
- Model:

$$\ln(HEAD_i / LIMB_i) = \alpha + \beta(HELMET_i) + \varepsilon_i, \qquad (1)$$

where ε_i are assumed to be i.i.d. normal random variables

- Estimated 24%, 32% and 28% reduction in head injury due to the helmet law for primary, secondary and adult cyclists in non-motor vehicle crashes
- Estimated 20% reduction overall for motor vehicle crashes

Robinson (2001)

Robinson (2001)

- Suggested those effects were "an artefact caused by failure to fit time trend trends" in model (1)
- To illustrate the idea, some "simulated data" were created: ratio of head to limb injuries falls by 0.1 per year
- Model:

$$HEAD_i / LIMB_i = \alpha + \beta(HELMET_i) + \varepsilon_i,$$

- A highly significant estimate for β was obtained
  - ➢ "spurious" because the data contains no effect of helmet wearing, only a linear trend!

UNSW

TARS Research

Robinson (2001)

Robinson (2001)

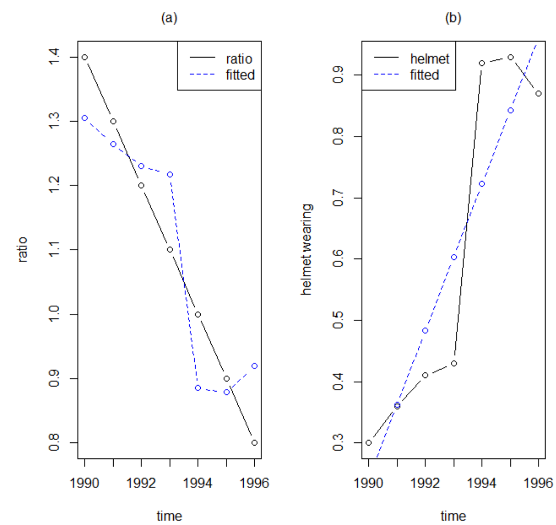
- Result not really "spurious" because HELMET is highly correlated with TIME (R=0.90)
- We regressed HELMET on TIME and slope estimate is highly significant (p=0.0056)
- Even the "simulated" data contains no effect of helmet wearing, HELMET is significant in predicting the ratio of head to limb injuries since itself can be predicted by TIME
- In fact, corr(HELMET, "simulated" data)=corr(HELMET, TIME)
  - ➢ Use the time dependent component of HELMET
  - ➢ Remaining effect of HELMET does not improve the model

UNSW

TARS Research

Robinson (2001)

Robinson (2001)



Povey et al. (2001) & Robinson (2001)

What are the problems?

- Data are not "simulated" → not generated from any model with random errors
- Robinson used head to limb ratio for primary and secondary school children to estimate trend → what about just a linear time trend?
- Given the data in Robinson (2001) were correct, results in Povey et al. (1999) could not be reproduced
- Neither of these studies have checked for model assumptions

$$\varepsilon_i \overset{i.i.d}{\sim} N(0,\sigma^2)$$

Povey et al. (2001) & Robinson (2001)

Time trend

- We fit a simple linear regression to the ratio of adult head to limb injuries using just a linear time variable
- Model:

$$\ln(HEAD_i / LIMB_i) = \alpha + \delta(TIME_i) + \varepsilon_i, \qquad (2)$$

- Compute fitted values and compare with Povey et al. and Robinson
- Criterion: mean squared error

$$(MSE) = \frac{\text{sum}(\text{predicted} - \text{actual})^2}{\text{number of cases}}$$

ACRS Conference, 2013 11

Povey et al. (2001) & Robinson (2001)

Time trend

Year	(R=HI/L)	Prediction of R			Helmet wearing (%)
		Povey et al.	**Robinson**	**Time trend**	
1990	1.40	1.17	1.25	1.28	30
1991	1.09	1.13	1.15	1.17	36
1992	1.07	1.10	1.16	1.07	41
1993	0.94	1.09	1.00	0.98	43
1994	0.86	0.81	0.80	0.89	92
1995	0.83	0.80	0.85	0.81	93
1996	0.77	0.83	0.75	0.74	87
MSE		**0.0113**	**0.0056**	**0.0031**	
Changes					
1990-1993	-0.45	-0.08	-0.25	-0.30	
1993-1995	-0.11	-0.29	-0.15	-0.17	

ACRS Conference, 2013 12

HELMET and Time trend

- Based on model (1) of Povey et al., we examine the effect of adding a linear time trend
- Model:

$$\ln(HEAD_i / LIMB_i) = \alpha + \beta(HELMET_i) + \gamma(TIME_i) + \varepsilon_i, \quad (3)$$

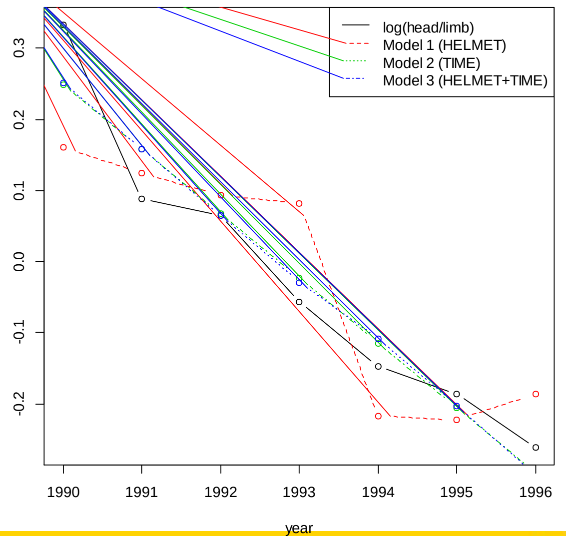
- Criteria
  - Adjusted R^2
  - Residual standard error
  - Akaike Information Criterion (AIC)

Parameter estimates

	M1 (HELMET)	M2 (TIME)	M3(HELMET+TIME)
Estimate of α	0.34 (0.07, 0.61)	0.34 (0.22, 0.46)	0.37 (0.17, 0.50)
p-value	0.0221	0.0008	0.0046
Estimate of β	-0.61 (-1.02, -0.20)	-	0.03 (-0.54, 0.60)
p-value	0.0123	-	0.8858
Estimate of γ	-	-	-0.09 (-0.17, -0.02)
p-value	-	-	0.0256
Estimate of δ	-	-0.09 (-0.12, -0.06)	-
p-value	-	0.0004	-
Adjusted R^2	0.6943	0.9233	0.9046
Residual S.E.	0.1122	0.0562	0.0627
AIC	-7.12	-16.79	-14.83

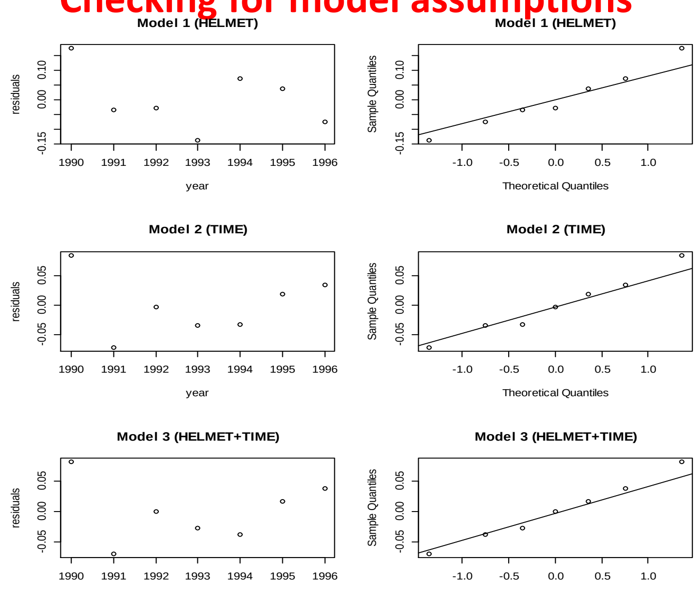
Povey et al. (2001) & Robinson (2001)

Fitted values under three models



Povey et al. (2001) & Robinson (2001)

Checking for model assumptions



Povey et al. (2001) & Robinson (2001)

Discussion

- Not appropriate to have both HELMET and TIME in the model (model 3) → multicollinearity
- Model 2 with only a linear time trend provides the best fit to the data
  - model assumption may not be satisfied
  - answer to the research hypothesis?
- Time trend
  - Need to account for it in the model?
  - Or more about serial correlation? → model assumption
  - More involved time series modelling techniques (Commandeur et al. 2012)
- Lack of data
  - Yearly data, n=7
  - Degrees of freedom = 5, 5, 4 for model 1, 2, 3 respectively

Clarke (2012)

Clarke (2012)

- Claim: a 51% drop in the average number of hours cycled per person from 1989-1990 period to 2003-2006 period, completely attributable to the helmet law
- Problems:
  - cycling rates near the introduction of the helmet law are more likely to be influenced by law → no data are presented around the date of the helmet law
  - comparing two numbers on either side of the law does not account for background trend → decline in ridership began long before the helmet law (Tin Tin, 2009)
- Claim: compared with 1988-1991, cyclists had a 20% higher accident rate by 2003-2007
- Problem:
  - Compare pre-law data to 1996-1998, there is 17% drop in cyclist injuries overall and 53% drop in serious injuries with AIS ≥ 3

Clarke (2012)

Clarke (2012)

- Claim: cyclist safety, compared to pedestrians, has reduced appreciably from 24% (1989-1993) to 49% (2006-2009) (cyclist deaths)
- Problems:
  - there is a 23% decline in cyclist fatalities in the immediate three years post-law (1994-1996)
  - Possible confounding factors?
- How useful are the data presented?
  - Helmet laws aim to increase helmet wearing to mitigate bicycle head injuries
  - Fatalities and injury counts are for all bicycle related injuries, cannot be used to estimate head injuries before and after the helmet law
- How reasonable is the conclusion: mandatory helmet law halved the number of cyclists and contributed to 53 deaths each year?

Scuffham and Langley (1997)

Scuffham and Langley (1997)

- Examine serious injury trends for three age groups of cyclists between 1980 and 1992; 2 years before the helmet law
- A Poisson regression model is used for number of injured cyclists with a head injury
  - Total number of cyclists admitted used as offset
  - Covariates: admission policy variable, helmet wearing, time
- Results:
  - No significant difference in downward trend between age groups
  - Only significant variable in the model is time
  - Downward trend in head injuries was due to time trend and was independent of helmet wearing
- Potential problems:
  - Multicollinearity: helmet wearing and time are highly correlated
  - Checking for model assumptions?

Scuffham et al. (2000)

Scuffham et al. (2000)

- Used a similar model as the one in Scuffham and Langley (1997)
  - Negative binomial was used instead of the Poisson distribution
  - Data between 1988 and 1996
- No linear temporal trend included in the model
  - Addition of a time-trend variable caused the helmet wearing to be insignificant
  - A time trend variable “swamped” the “real effect”
- Results:
  - A negative and significant estimate for helmet wearing variable
  - Helmet law has been an effective road safety intervention that has lead to a 19% reduction in head injury over the first 3 years
- Potential problems:
  - Substantial seasonal pattern not accounted for → dummy variables for seasonal patterns/X11 methods

ACRS Conference, 2013 21

Conclusions

Conclusions

- Povey et al. (1999) :
  - Omitted checking for model assumptions
  - Results and conclusions are valid after we checked the assumptions
  - Proper data and methods to test a well-defined hypothesis
- Robinson (2001):
  - “simulation” study is demonstrably flawed
  - Omitted checking for model assumptions
- Clarke (2012)
  - No statistical inference, purely descriptive
  - Due to weakness in the analysis, conclusion (MHL in NZ halved the number of cyclists and contributed to 53 deaths each year) is highly questionable if not misleading

ACRS Conference, 2013 22

Recommendations

Recommendations

- Do you have enough data?
  - Use monthly data instead of highly aggregated yearly data if possible
- Are your data useful for answering your research question?
- Do you need to include temporal trend?
  - Checking for serial correlation
  - Is it appropriate to add a linear time component?
  - Given large amount of data, use specialised time-series models such as ARIMA
- Have you checked for model assumptions?
  - Results are erroneous/misleading if assumptions are not satisfied
- When adding two or more variables, have you checked their correlations?
  - Multicollinearity: estimates less precise/results are misleading

Thank you!