TRAFFIC INJURY RESEARCH FOUNDATION

AN EVALUATION OF GRADUATED DRIVER LICENSING PROGRAMS IN NORTH AMERICA

AN ANALYSIS OF RELATIVE FATALITY RISKS OF 16, 17, 18 AND 19 YEAR OLD DRIVERS USING A META-ANALYTIC APPROACH

TIRF

The knowledge source for safe driving

AN EVALUATION OF GRADUATED DRIVER LICENSING PROGRAMS IN NORTH AMERICA

An Analysis of Relative Fatality Risks of 16, 17, 18 and 19 Year Old Drivers Using a Meta-Analytic Approach

Ward Vanlaar*, Dan Mayhew*, Kyla Marcoux*, Geert Wets**, Tom Brijs**, Jean Shope***

* Traffic Injury Research Foundation (TIRF)
171 Nepean St., Suite 200
Ottawa, ON, K2P 0B4

** Transportation Research Institute
Hasselt University – Campus Diepenbeek
Wetenschapspark 5 bus 6
BE-3590 Diepenbeek

*** University of Michigan Transportation Research Institute (UMTRI)
2901 Baxter Rd.
Ann Arbor, MI 48109-2150 – USA

Traffic Injury Research Foundation

ISBN: 978-0-920071-84-7

ACKNOWLEDGEMENTS

A draft of this report was circulated for critical review to Jim Hedlund, Highway Safety North; Paul Boase, Transport Canada; and, Allan Williams. The authors extend their appreciation to each of them for their insightful comments, perspectives and suggestions.

Herb Simpson, Traffic Injury Research Foundation provided additional feedback to an earlier draft of this report. Additional project support was provided by staff at the Traffic Injury Research Foundation, Sara Oglestone, Gisèle Perron, and Amanda Johnson. Their contribution to the completion of this work is greatly appreciated.

Funding for this project was graciously provided by Transport Canada and the Transportation Research Institute at Hasselt University, Belgium.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Most jurisdictions in North America have some version of graduated driver licensing (GDL). GDL programs attempt to provide a more protective environment for novice drivers, typically by lengthening the learning process and imposing a set of restrictions aimed at reducing their risk of collision. To achieve this, most GDL programs are multi-staged and include a learner's stage and an intermediate stage before graduation to a full license. A sound body of evidence documenting the effectiveness of GDL programs in reducing collisions, fatalities and injuries among novice drivers is available. However, information about the relative importance of individual components of GDL is lacking. Despite the available literature it is still not known which GDL features contribute most to collision reduction and how exactly this is achieved. Consequently, it is difficult to identify how a GDL program should be best designed or improved.

The objectives of this study are to calculate a summary statistic of GDL effectiveness, to identify the most effective components of GDL programs, and to help understand how GDL components achieve their effect by applying a meta-analytic approach.

Data from 46 American States, the District of Columbia and 11 Canadian jurisdictions are used and were obtained from the Fatality Analysis Reporting System (FARS) for the U.S. and from Transport Canada's Traffic Accident Information Database (TRAID) for Canada. The timeframe of this evaluation is 1992 through 2006, inclusive. Relative fatality risks and their and 95% confidence intervals were calculated using fatality counts and population data for target and comparison groups, both in a pre-implementation and post-implementation period in each jurisdiction. The target groups were 16, 17, 18 and 19 year old drivers. The comparison groups were 25-54 year old drivers. The relative fatality risks of all jurisdictions were summarized using the random effects DerSimonian and Laird model. Cumulative meta-analyses and meta-regression using Restricted Maximum Likelihood (REML) and Markov Chain Monte Carlo (MCMC) Gibbs sampling were also conducted.

Strong evidence in support of GDL was found. GDL had a positive and significant impact on the relative fatality risk of 16 year old drivers (reduction of 19.1%). However, no such summary effects were found for 17, 18 and 19 year old drivers.

Significant effects were found for meta-regression models with 16, 18 and 19 year old drivers. Two variables were significant for 16 year old drivers, namely whether there are restrictions on passengers in the intermediate stage (in jurisdictions with passenger restrictions in the intermediate stage, the relative fatality risk of 16 year old drivers decreases) and whether passenger restrictions are lifted or not in the intermediate stage if passengers are family members (lifting the passenger limit in the intermediate stage if passengers are immediate family members leads to an increase in the relative fatality risk of 16 year old drivers).

One variable was significant with 18 year old drivers, more precisely mandatory driver education in the learner stage (the relative fatality risk of 18 year old drivers decreases in jurisdictions where driver education is mandatory in the learner stage).

Finally, five variables were significant with 19 year old drivers. These variables are length of night restriction in the learner stage (an increase in length of the night restriction in the learner stage leads to an increase in the relative fatality risk of 19 year old drivers), country (the relative fatality risk of 19 year old Canadian drivers is higher than that of 19 year old drivers in the U.S.), lifting night restrictions in the intermediate stage for work purposes (the relative fatality risk of 19 year old drivers increases in jurisdictions where night restrictions are lifted for work purposes in the intermediate stage), exit test in the intermediate stage (the relative fatality risk of 19 year old drivers decreases in jurisdictions that require an exit test to graduate from the intermediate stage), and mandatory driver education in the intermediate stage (there is an increase in the relative fatality risk of 19 year old drivers in jurisdictions with mandatory driver education in the intermediate stage). These results are described, interpreted and further discussed along with the limitations of this study.

In conclusion, despite limitations of the study design, some previously established findings have been confirmed and some interesting and intriguing new findings emerged from these analyses. Recommendations for follow-up research are formulated based on the conclusions of this report.

INTRODUCTION

Background

Graduated driver licensing (GDL) programs attempt to provide a more protective environment for novice drivers, typically by lengthening the learning process and imposing a set of restrictions aimed at reducing their risk of collision. To achieve this, most GDL programs are multi-staged and include a learner's stage and an intermediate stage before graduation to a full license.

Despite the "somewhat torturous journey that graduated licensing has experienced in achieving acceptance among the public and policy-makers" (Simpson 2003: p. 25), GDL programs are now commonplace. Indeed, most jurisdictions in Canada and the United States have some version of GDL. Furthermore, a consensus exists in the research community about the effectiveness of GDL programs and today they are widely accepted as an effective safety measure: "The systems that have been evaluated have been found to be very effective in reducing crashes and injuries, and public acceptance is high. This in and of itself provides the compelling case for graduated licensing." (Williams 2003: p. 3).

A sound body of evidence documenting the positive influence of GDL programs on collisions among novice drivers is available. In addition to the more than two dozen evaluation studies of specific GDL programs (see Mayhew et al. 2005 for a descriptive review of these evaluations; see also Hedlund et al. 2006 for an update on research published since 2005 and research in progress; and Shope 2007 for the most recent review of the GDL evaluation literature), two systematic reviews of this literature have been carried out (Foss and Evenson 1999; Hartling et al. 2005). Collectively, these studies and their reviews clearly establish GDL as an effective safety measure.

The authors of a recently published evaluation study that distinguished between different kinds of GDL programs came to the same conclusion (Morrisey et al. 2006; see also Dee et al. 2005). This study, however, was limited geographically in that it only analyzed data from the United States. A more recent study (Baker et al. 2006) was carried out to determine which types of GDL programs are associated with reductions in fatal crashes involving 16-year-old drivers — evaluations have found a wide range in GDL effectiveness so this particular study sought to determine which types of programs had the greatest impact. This study is also limited geographically to the U.S. and only investigated one age cohort. Furthermore, while the results of this study allow distinguishing between more effective programs (those with at least five GDL components) and less effective programs (those with less than five GDL components), the authors did not estimate the relative importance of individual components of GDL programs — components such as nighttime driving restrictions, supervision or passenger restrictions. Some of the same limitations apply to a more recent study of GDL effectiveness, both in terms of fatal and injury collision involvements, by the same authors (Baker et al., 2007).

Finally, Williams (2007) conducted a literature review to analyze the available evidence regarding GDL components and concluded that extended learner periods, nighttime restrictions and passenger restrictions have contributed to crash reductions, but that there is more to learn about GDL and its components.

Therefore, despite the available literature it is still not known which GDL features contribute most to collision reduction and how exactly this is achieved. Consequently, it is difficult to identify how a GDL program should be best designed or improved. As a result, a priority research need is to identify "effects of specific GDL components and provisions" (Hedlund et al. 2003: p. 109), a need that had already emerged from the Symposium on Graduated Driver Licensing in Chatham, MA on November 5 – 7, 2002.

Objectives and Rationale

The objectives of this study are to calculate a summary statistic of GDL effectiveness, to identify the most effective components of GDL programs, and to help understand how GDL components achieve their effect by applying a meta-analytic approach.

In their review of the existing literature, Foss and Evenson (1999) were unable to conduct a meta-analysis of GDL evaluation studies and to draw reliable conclusions about the impact of different features of GDL programs on crash numbers, probably because the number of studies available to them at that time was rather small (seven). At best, Foss and Evenson (1999) had to compare outcomes from different studies, using different methodologies. Without controlling for these differences, it is difficult to compare the results of program effectiveness, which, in turn, makes it difficult to determine which features of programs are most effective.

The ability to address this challenge is an advantage of meta-analysis that controls for the different methodologies by comparing the different studies available "on the same scale", i.e., standardizing. Results of a meta-analysis regarding differences in effectiveness will, therefore, be more valid because they truly reflect differences in programs and program features by accounting for differences in evaluation methodologies (see Elvik 2005a and 2005b).

A more recent effort by Hartling et al. (2005) to systematically review the GDL literature used the well-acknowledged Cochrane methodology (www.cochrane.org). Even though the number of evaluation studies included in their review was higher than in the Foss and Evenson study (1999), the authors decided it was not appropriate to pool the results of the evaluation studies and perform a meta-analysis, "due to statistical heterogeneity and differences among studies with respect to study quality and design, program quality and design, definition of outcomes, baseline rates, and data reported" (Hartling et al. 2005: p. 5).

Their view is shared by the authors of this report and led to the investigation of the feasibility of obtaining raw data for each of the evaluated studies. Using such raw data to conduct a meta-analysis is considered to be "the 'yardstick' against which other forms of systematic review should be measured" (Clarke and Stewart 2007: p. 110) because it becomes possible to standardize the outcome measure to the fullest. In other words, by using raw data to calculate a standardized outcome measure it is no longer necessary to use the different kinds of outcome measures that are being reported in evaluation studies and that are not necessarily suitable to be summarized into one statistic. It was therefore decided to use counts of fatalities per age cohort and by state from the Fatality Analysis Reporting System (FARS) for the U.S. and comparable data for Canadian jurisdictions, contained in Transport Canada's Traffic Accident Information Database (TRAID), to calculate a standardized outcome measure. Once it was decided to use these raw data, it was no longer desirable to limit the evaluation to those jurisdictions for which an evaluation study was available; hence the present evaluation includes all U.S. and Canadian jurisdictions that have a GDL program in place.

While the choice to use raw data to calculate a standardized outcome measure for each of the jurisdictions included in this evaluation precluded any difficulties with respect to methodological heterogeneity among different studies, there was still the considerable heterogeneity with respect to programs and their features to deal with; see for example the on-line inventories of GDL programs on the website of the Insurance Institute for Highway Safety (IIHS) and the Traffic Injury Research Foundation (TIRF), respectively at www.iihs.org/laws/graduatedLicenseIntro.aspx and www.trafficinjuryresearch./yndrc/default.asp

However, sophisticated meta-analysis techniques are available today to explicitly model diversity, more precisely random meta-regression analysis using multilevel techniques and Bayesian statistics in multilevel modeling (Goldstein 2003). "The major advantage of using multilevel analysis instead of classical meta-analysis methods is flexibility. In multilevel analysis, it is simple to include study characteristics as explanatory variables in the model. If we have hypotheses about study characteristics that influence the outcomes, we can code these and include them on a priori grounds in the analysis. Alternatively, after we have concluded that the study outcomes are heterogeneous, we can explore the available study variables in an attempt to explain the heterogeneity." (Hox and de Leeuw 2003: p. 92).

These technical advantages and the flexibility of using 'multilevel' or 'random effects' meta-regression rather than 'fixed effect' meta-analysis can also be expressed in terms of a conceptual advantage. The fixed effect model assumes that variation between programs or heterogeneity is exclusively due to random variation and, therefore, if the data available about each program were infinitely large, the results would be identical (Egger and Smith 2007). Since fixed effect analysis considers a common effect across programs (Smith et al. 1995), this would be equivalent to assuming that each program is equally effective and that differences in effectiveness are really only due to random fluctuations, rather than to, for example, GDL program features. In other words, using a fixed effect model really implies an *a priori* choice that no differences in effectiveness between programs exist, regardless of the different composite features of each program.

A random effects model, on the other hand, "assumes a different underlying effect for each study and takes this into consideration as an additional source of variation" (Egger and Smith 2007: p. 35), which is "mathematically equivalent to assuming these effects are drawn from some population" (Smith et al. 1995: p. 2685). Since this study is predicated on the assumption that GDL programs can be improved through the enhancement of their different features, it seems preferable to at least test for the presence of heterogeneity and to use a random effects model to analyze the data accordingly, rather than to simply model the data using a fixed effects model. This approach was adopted in the current study.

METHOD

Study Population and Data

The analyses in this study use data from 46 American States, the District of Columbia and 11 Canadian jurisdictions (see the appendices for a list of jurisdictions). The timeframe is 1992 through 2006, inclusive. Jurisdictions were excluded from the analyses only if post data at the level of jurisdictions were not available, e.g., because the implementation of the GDL program occurred too recently (note that at least two years of data post-implementation were needed to calculate the outcome measure as explained below). These jurisdictions include Arizona (implementation in July 2007), Hawaii (implementation in September 2006), Montana (implementation in July 2006), Kansas (pursuing legislation in 2008) and Canada's North West Territories (implementation in 2005) and Nunavut (no GDL program in place).

If legislative changes to the initial GDL program were passed and took effect in a particular jurisdiction, this jurisdiction was included at least twice, i.e., once to reflect the original implementation of the program and once to reflect the legislative change. For example, the GDL program in British Columbia (BC) was implemented in 1998 and improved in 2003, so two data points for BC were included in the master data file, and the corresponding outcome measure and independent variables for the jurisdiction were measured accordingly. As such, several jurisdictions have been included more than once (see appendices for a list of jurisdictions and the corresponding dates of implementation or legislative changes) so the master database contains 78 data points rather than 58 (46 States, Washington, DC and 11 Canadian jurisdictions).

Fatality rates were calculated separately for 16, 17, 18 and 19 year old drivers using counts of fatalities per age cohort and by jurisdiction from FARS data for U.S. jurisdictions and TRAID data for Canadian jurisdictions. Population data for each jurisdiction were obtained from the U.S. Census Bureau for U.S. jurisdictions (see http://www.census.gov/popest/archives/1990s/st_age_sex.html for estimates for 1992-1999 and http://www.census.gov/popest/datasets.html for estimates for 2000-2007) and from Statistics Canada's 2007 Demographic Estimates Compendium for Canadian jurisdictions (see Statistics Canada 2007 for a detailed description of the methodology and the quality of Canadian population estimates). Both IIHS's and TIRF's websites were used to obtain descriptions of the GDL programs for each of the included jurisdictions in the analyses. This information was then coded into a set of 23 independent variables. These variables are described in more detail in Table 1 and were included as covariates in the meta-regression analyses.

Outcome Measure

The outcome measure in this study, i.e., the dependent variable, was calculated as described in Altman and Bland (2003) and applied by Ulmer et al. (2000), Mayhew et al. (2001, 2002), Foss et al. (2001) and Shope et al. (2001). More precisely, for each jurisdiction eight numbers were obtained:

> the number of fatalities in the post period for the target group (a);
> the population in the post period for the target group (b);
> the number of fatalities in the pre period for the target group (c);
> the population in the pre period for the target group (d);
> the number of fatalities in the post period for the comparison group (e);
> the population in the post period for the comparison group (f);
> the number of fatalities in the pre period for the comparison group (g);
> the population in the pre period for the comparison group (h).

The target groups were 16, 17, 18 and 19 year old drivers. In all cases, the comparison group was 25-54 year old drivers, who are assumed to be largely unaffected by the GDL program in a jurisdiction.

The post period was defined as a period of 12 months, starting one year after the implementation of the GDL program and ending two years after implementation. The pre period was defined as a period of 12 months, starting two years before the implementation of the GDL program and ending one year before the implementation. If the program was implemented more recently, the post and pre periods were adjusted accordingly – for example GDL was implemented in mid September 2005 in Wyoming so only 3.5 months of post information was available (mid September 2006 to 31 December 2006) and only 3.5 months of pre information was used (June 2004 to mid September 2004).

The reason why such timeframes have been chosen is because it has been shown that the implementation of a GDL program can disrupt the normal licensing patterns. For example, normal licensing patterns can be disrupted before GDL implementation due to novices trying to avoid the change and after GDL implementation due the time needed for drivers to meet new requirements, as well as to progress through the different GDL stages. This may also affect crash patterns (see e.g., Mayhew et al. 1999, 2001).

The available information was then summarized into fatality rates for the target group ((a/b)/(c/d)) and the comparison group ((e/f)/(g/h)) and then into a fatality ratio (or relative fatality risk) by dividing each fatality rate for each target group (e.g., the fatality rate for drivers aged 16) by the fatality rate for the comparison group (i.e., the fatality rate for drivers aged 25-54). Note that only drivers who died in a fatal crash were counted for the numerators (a, c, e and g), while population numbers pertaining to the entire population and not just drivers were used for the denominators (b, d, f and h). This process of obtaining the fatality rates per age cohort and by jurisdiction and calculating the different relative fatality risks (four in total) for each of the 'jurisdictions' (78 in total) was automated in Stata, release 10 (StataCorp. 2007) with a variety of do-files and automatic do-files.

Using such a relative fatality risk as an outcome measure standardizes the fatalities of a target group to the population of that group as well as to the fatality rate of the comparison group. A ratio of less than one indicates a positive impact of GDL on the fatality risk of young drivers (i.e., a decrease from pre to post period) relative to the comparison group of older drivers in that jurisdiction. The standard error (s.e.) of this measure is used to calculate whether this positive impact is significantly different from the trend in the comparison group or not.

The different analyses used to summarize the data (described below) all assume that the outcome measure and its s.e. are measured on the log-scale, so they have been transformed for further analyses and rescaled using the exponential function for the interpretation of the effect of the independent variables on the outcome measure.

Graduated Driver Licensing Programs

A description of the different North-American GDL programs included in this analysis can be found on IIHS's and TIRF's websites. Not surprisingly, there is a lot of variation among the different programs making it particularly challenging to formally describe them in the form of variables that can be included in a meta-regression analysis to investigate potential sources of heterogeneity in the outcome measure. Nevertheless, about two dozen variables were used in this study to capture such differences that can help explain which GDL features are more effective than others, and why. Table 1 contains a description of each of the independent variables used in this study.

Table 1: Description of the independent variables included in the meta-regression analysis (variable label; categories and frequencies for categorical variables and range and mean for numerical variables)

Variable label	Categories/range (frequencies/mean)
effective date of implementation or legislative change	1992-2005
IIHS's rating of the quality of the GDL program	good (28), fair (18),
	marginal (12), poor
minimum length in months of mandatory holding period learner stage	0-12 (6.0)
maximum length in months of mandatory holding period learner stage	0-48 (6.8)
minimum # of hours of supervisory driving required in learner stage	0-60 (22.0)
conditions under which supervisory driving occurs	0=no mandatory hours at night (43)
	1=mandatory hours at night (35)
length of night restriction in hours in learner stage	0-10 (1.3)
night restrictions lifted if supervised	0=no (71)
	1=yes (3)
restriction on passengers in learner stage	0=no restrictions (67)
	1=restrictions (11)
passenger limit lifted in learner stage if passengers are immediate family members	0=no (76)
	1=yes (2)
pasenger restriction lifted for family in learner stage if driver is accompanied by a licensed instructor and driver is in driver education	0=no (77)
	1=yes (1)
minimum entry age for learner stage	14-16 (15.3)
reduction in # of months of mandatory holding period for taking driver education (time discount)	0-8 (0.6)
country	0=US (47)
	1=Canada (12)
driver education requirements in learner stage	0=no requirements (35)
	1=driver education mandatory (17)
	2=time discount if driver ed. (8)
length of night restriction in hours in intermediate stage	0-10 (4.1)
night restrictions lifted for work purpose in intermediate stage	0=no (74)
	1=yes (2)
restriction on passengers in intermediate stage	0=no (31)
	1=yes (47)
passenger limit lifted in intermediate stage if accompanied by a qualified supervisor	0=no (74)
	1=yes (3)
passenger limit lifted in intermediate stage if passengers are immediate family members	0=no (41)
	1=yes (36)
minimum entry age for intermediate stage	14.5-17 (16.1)
driver education requirements in intermediate stage	0=voluntary (54)
	1=mandatory (2)
exit test required to graduate from intermediate stage to full stage	0=no (50)
	1=yes (7)

Model and Data Analysis

The data in this study were analyzed using three different approaches. First, in an exploratory phase, a summary effect was calculated for each target group and a test for heterogeneity conducted using the random effects DerSimonian and Laird model, as described in Deeks et al. 2007. This analysis was carried out in Stata, release 10 (StataCorp. 2007), using the metan-command. Furthermore, a cumulative meta-analysis was carried out as well in Stata, using the metacum-command and the same random effects model. The summary effect for each age cohort is based on the complete database of 78 'jurisdictions'.

In a second step, a meta-regression was carried out in Stata (using the metareg-command; see Harbord and Higgins 2008) to investigate the relative influence of the different independent variables on the outcome measure, using Restricted Maximum Likelihood (REML); see Sterne et al. 2007a.

The model used to summarize the data, expressed as a multilevel model is shown below (see Hox and de Leeuw 2003):

$$d_j = \gamma_0 + \gamma_1 Z_{1j} + \gamma_2 Z_{2j} + \ldots + \gamma_p Z_{pj} + u_j + e_j \qquad \text{equation 1}$$

In equation 1 d_j denotes the observed outcome of jurisdiction j (j=1,...,J); γ_0 is the intercept, i.e., the estimate for the mean outcome measure across all jurisdictions, or the summary effect; $\gamma_1 \ldots \gamma_p$ are the regression coefficients; u_j is the residual error term at the level of jurisdictions and its variance, σ_u^2, represents the true variation between jurisdictions and is assumed to have a normal distribution; and e_j is the residual error term representing the difference between the jurisdictions that is the result of sampling variation, which is determined entirely by the within-jurisdiction variation and sample size, and is assumed to be known from the jurisdictions.

Finally, a full Bayesian analysis (Smith et al. 1995) was conducted in the MLwiN statistical package, using Markov Chain Monte Carlo (MCMC) Gibbs sampling (see Browne 2004). The meta-regression using the REML estimation procedure in the second step and the MCMC Gibbs sampling in this third step are both based on a sample of 48 'jurisdictions' (rather than 78), due to missing data for the variables included as covariates in these models. The results obtained from the random meta-regression in the second step are used as starting values for each of the parameters that need to be estimated with the MCMC Gibbs sampling procedure.

The length of each MCMC chain was set at 50,000 iterations with a burn-in period of 1,000 iterations. Each of the models was estimated four times with different random number seeds to ensure the results are stable when using different starting values (the results of these different models were found to be similar; see appendices 5, 6 and 7). Diagnostics for each parameter were obtained as well to compare different models. These include graphs of Gibbs sampling traces to check for autocorrelation in these traces (see Browne 2004), the Raftery-Lewis diagnostic (Raftery and Lewis 1992) and the Brooks-Draper diagnostic (see Browne 2004) to assess the required length of the MCMC chains and the Deviance Information Criterion (DIC) according to Spiegelhalter et al. (2002), which is a generalization of Akaike's Information Criterion (AIC) to test the complexity of the different models and their goodness of fit. Finally, the different jurisdictions were also ranked according to their effectiveness, although this turned out to be meaningless because the credibility intervals (confidence intervals in Bayesian terminology are called credibility intervals) for each jurisdiction obtained with the MCMC Gibbs sampling all overlapped.

For a discussion about the advantages and disadvantages of a Bayesian approach used in step three of the analyses versus the more classical ‘frequentist’ approach, used in steps one and two of the analyses, see Carlin and Louis (2000). For a formal description of the Bayesian model used in this report, see Browne (2004) and Smith et al. (1995). For a technical description about how to use MLwiN to run these analyses, see Lambert and Abrams (1995) and Turner et al. (1999).

RESULTS

Summary Effect

Summary effect for 16 year old drivers

Appendix 1 contains a table (entitled "Table 1.1: Random Effects Meta-Analysis") showing the outcome measure for 16 year old drivers (ES), i.e., the relative fatality risk (see subsection entitled 'Outcome Measure'), its 95% confidence interval (95%-CI) and its weight for each jurisdiction included in the evaluation, as well as a pooled summary effect for all jurisdictions according to the random effects model. The weight is derived from the variance of each jurisdiction, which means that smaller jurisdictions with more variance will contribute less to the pooled estimate, while larger jurisdictions with less variance will contribute more.

As can be seen, the pooled summary effect for 16 year old drivers is 0.809, with a 95%-CI of (0.714-0.917). The null hypothesis of no effect is rejected (z=3.32, p=0.001). This means the evaluation provides strong evidence in support of GDL because the outcome measure is significantly smaller than one. More precisely, GDL has had a positive and significant impact on fatalities among 16 year old drivers — a decrease in the relative fatality risk of 19.1% ((1-0.809)*100) — when adjusting for a group of older drivers who are assumed not to have been affected by the implementation of GDL.

The results of the test of homogeneity are also displayed in Table 1.1 in Appendix 1 and indicate that the null hypothesis of homogeneity cannot be rejected (chi-square=81.17; d.f.=77; p=0.351). In other words, there is no evidence for heterogeneity (i.e., differences in the outcome measure that can be accounted for by differences among GDL programs) according to this test. However, the test of homogeneity has low power (Deeks et al. 2007; Thompson 2007), which means the possibility of a type II error (false negative, or, deciding there is no heterogeneity between jurisdictions while in reality there is) must always be considered. As such, "it is often more useful to quantify heterogeneity than to test for it" (Harbord and Higgins 2008: p. 499). The I-squared measure quantifies heterogeneity (Higgins et al. 2003) and shows that 5.1% of the variance in the outcome measure is due to heterogeneity. Potential sources of heterogeneity will be investigated using meta-regression analysis techniques in the next section.

Appendix 1 also contains a forest plot (see Figure 1.1). In a forest plot the contribution of each jurisdiction (its weight) is represented by the area of a box whose centre corresponds to the size of the effect estimated from that jurisdiction. The 95%-CI for the effect from each jurisdiction is also shown and the summary effect is represented by the middle of a diamond whose left and right extremes represent the corresponding confidence interval (Sterne et al. 2007a). This forest plot displays the same results from the table in this appendix, but in a different format.

Finally, a figure from a cumulative meta-analysis is inserted in Appendix 1 as well (see Figure 1.2). This figure shows that there has been consistent albeit not significant evidence in support of GDL, since approximately 1996, around the time when Virginia (VA) implemented its program. The pooled summary effect was smaller than one from then on (the horizontal lines are 95%-CIs, the ovals are the point estimates, and the vertical line corresponds to the overall summary effect). Note that findings from this cumulative analysis do not say anything about individual results from Virginia or any other jurisdiction *per se*. Since 2000, with the implementation of Oregon's program, the overall effect also became significantly smaller than one (95%-CI: 0.710-0.994) and stayed significantly smaller than one. Since then, the effect grew gradually stronger — albeit only slightly stronger — in favor of GDL, as can be seen both from the summary effect, which is more removed from one and from its 95%-CI, which has become narrower.

Summary effect for 17 year old drivers

Appendix 2 contains a table (Table 2.1) showing the outcome measure for 17 year old drivers. This table uses the same format as the one for 16 year old drivers.

As can be seen, the pooled summary effect for 17 year old drivers is 1.001. However, the 95%-CI for this age cohort (0.906-1.105) contains one, meaning the relative fatality risk is not significantly different from one, and the true effect could be greater or smaller than one. The null hypothesis of no effect is not rejected (z=0.02, p=988). This means the evaluation provides no evidence in support of, or against GDL based on data from 17 year old drivers. In other words, the available data do not support the contention that GDL has had a positive and significant impact on the number of fatalities among 17 year old drivers when adjusting for a group of older drivers who are assumed not to have been affected by the implementation of GDL.

The results of the test of homogeneity are also displayed in Table 2.1 in Appendix 2 and indicate that the null hypothesis of homogeneity cannot be rejected (chi-square=53.16; d.f.=77; p=0.983). While there is still the possibility of a type II error occurring, the I-squared measure now shows that none of the variance in the outcome measure is due to heterogeneity. This will have to be kept in mind when further investigating potential sources of heterogeneity — it is unlikely that further investigation of this subpopulation's data will reveal significant relationships between independent variables (i.e., the GDL program features) and the dependent variable (i.e., the effectiveness of GDL in terms of the relative fatality risk) due to this lack of variability.

A forest plot and results from a cumulative meta-analysis are also available in Appendix 2 (Figures 2.1 and 2.2). As can be seen from the cumulative meta-analysis, the point estimate of the pooled summary effect has been very close — almost equal — to one since 1997 with the implementation of Quebec's program.

Summary effect for 18 year old drivers

Appendix 3 contains a table showing the outcome measure for 18 year old drivers (Table 3.1). This table again uses the same format as the one for 16 year old drivers.

The pooled summary effect is now equal to 1.083 and the 95%-CI for this age cohort (0.978-1.199) contains one. The null hypothesis of no effect is not rejected (z=1.53, p=0.126). This means the evaluation provides no evidence in support of, or against GDL based on data from 18 year old drivers. In other words, the available data do not support the contention that GDL has had a positive and significant impact on the number of fatalities among 18 year old drivers when adjusting for a group of older drivers who are assumed not to have been affected by the implementation of GDL.

The results of the test of homogeneity indicate that the null hypothesis of homogeneity cannot be rejected (chi-square=86.63; d.f.=77; p=0.212). According to the I-squared measure 11.1% of the variance in the outcome measure is due to heterogeneity.

A forest plot and results from a cumulative meta-analysis are also available in Appendix 3 (see Figures 3.1 and 3.2). As can be seen in Figure 3.2, a pattern comparable to the pattern of 17 year old drivers is apparent in that the pooled summary effect has never been significantly different from one, with the exception of one time, namely after the implementation of New York's program in 1992 when the pooled summary effect was significantly smaller than one.

Summary effect for 19 year old drivers

Appendix 4 contains a comparable table (Table 4.1) showing the outcome measure for 19 year old drivers.

The pooled summary effect is equal to 1.059 with a 95%-CI of (0.963-1.165). The null hypothesis of no effect is not rejected (z=1.19, p=0.235). This means the evaluation provides no significant evidence in support of, or against GDL based on data from 19 year old drivers. In other words, the available data do not support the contention that GDL has had a positive and significant impact on the number of fatalities among 19 year old drivers when adjusting for a group of older drivers who are assumed not to have been affected by the implementation of GDL.

The results of the test of homogeneity indicate that the null hypothesis of homogeneity cannot be rejected (chi-square=81.48; d.f.=77; p=0.342). According to the I-squared measure 5.5% of the variance in the outcome measure is due to heterogeneity.

A forest plot and results from a cumulative meta-analysis are also available in Appendix 4 (see Figures 4.1 and 4.2).

The Examination of Heterogeneity among Jurisdictions

In this section heterogeneity in the outcome measure is examined. Potential sources of heterogeneity have been formally described by a variety of variables (see Table 1) that have been included in a meta-regression analysis as covariates. As explained in the method section, two approaches were used to obtain the results. First, a random effects meta-regression analysis was run using REML. Then, MCMC Gibbs sampling was used to obtain full Bayesian estimates of the coefficients of the covariates. Results from both approaches are shown in Table 2 and compared in this section.

It warrants mentioning that only models for 16, 18 and 19 year old drivers are discussed in this section because variance in the outcome measure due to heterogeneity among jurisdictions was equal to zero for 17 year old drivers. The variance that was found for 16, 18 and 19 year old drivers on the other hand (5.1%, 11.1% and 5.5% respectively) may be low but makes it more likely to identify possible sources of heterogeneity (reasons why there is no variation among 17 year old drivers will be discussed later).

Meta-regression for 16 year old drivers

Table 2 provides an overview of the significant effects that were found for 16 year old drivers and compares the REML results with the MCMC results. These results come from a model that includes all covariates listed in Table 1. The full model can be found in Appendix 5 (see Table 5.1), as well as the MCMC diagnostics per significant parameter (see Figure 5.1).

Two effects were found to be significant according to the REML estimation and according to the MCMC Gibbs sampling (if a 95%-credibility interval does not contain zero, then the coefficient can be considered significant according to the MCMC Gibbs estimation). These effects are restrictions on passengers in the intermediate stage and whether passenger restrictions are lifted in the intermediate stage if passengers are immediate family members. As can be seen, coefficients from both estimation procedures are very similar — the coefficients according to the MCMC Gibbs sampling are somewhat smaller.

To assist with the interpretation of those parameters, the coefficients can be transformed using the exponential function. The exponentiated coefficient for the first variable then becomes 0.115 (REML p-value=0.014). The interpretation is as follows: in a jurisdiction with passenger restrictions in the intermediate stage, the relative fatality risk of 16 year old drivers decreases by a factor of 0.115 or 88.5% ((1-0.115)*100), compared to jurisdictions without such passenger restrictions.

The transformed result for the second variable is 8.281 (REML p-value=0.014). The interpretation of this coefficient implies that lifting the passenger limit in the intermediate stage if passengers are immediate family members leads to a 728.1% increase in the relative fatality risk of 16 year old drivers ((8.281-1)*100).

Table 2: Comparison of significant effects (on log-scale) according to REML and/or MCMC Gibbs for 16, 18 and 19 year old drivers

Variable	REML		MCMC Gibbs	
	Coefficient	*s.e. (p-value)*	*Coefficient*	*95%-Credibility interval*
16 year old drivers				
Passenger restriction in intermediate stage	-2.160	0.804 (0.014)	-2.102	-3.833;-0.364
No passenger restrictions in intermediate stage if passengers are family	2.114	0.794 (0.014)	2.011	0.237;3.762
18 year old drivers				
Driver education in learner stage mandatory	-0.423	0.189 (0.036)	-0.437	-0.876; -0.004
19 year old drivers				
Length night restriction in learner stage	0.104	0.047 (0.038)	0.102	0.020; 0.181
Country	2.587	1.374 (0.074)	2.543	0.391; 4.682
No night restriction in intermediate stage if work	3.953	2.089 (0.072)	3.966	0.675; 7.237
Driver education in intermediate stage mandatory	0.746	0.395 (0.073)	0.731	0.057; 1.394
Exit test to graduate from intermediate stage	-3.856	1.612 (0.026)	-3.803	-6.383; -1.195

The model was further investigated by checking for outliers and the relative influence of individual jurisdictions to find an explanation for these rather extreme effects. No single jurisdiction had an effect that was particularly greater than that of the other jurisdictions and no outliers were found (see Figure 5.2 in Appendix 5).

Also, it was found that 73.7% of the variance between jurisdictions (i.e., heterogeneity) is explained by the covariates in the model and that only 2.5% of the residual variation of the model is due to heterogeneity; the remaining 97.5% is due to within jurisdiction variation (see Table 5.1).

Finally, It was argued previously that a random effects model (be it one using REML or MCMC Gibbs) would be more appropriate for a variety of reasons (see section entitled "Objectives and Rationale"). This was formally tested using the Bayesian DIC. The Bayesian DIC was 123.72 for the random model while it was 135.79 for the fixed model. The lower value for the random model confirms such a model is indeed more appropriate than a fixed model because it has a better fit.

Meta-regression for 18 year old drivers

Table 2 also provides an overview of the results of the significant effect that was found for 18 year old drivers and compares both estimation procedures. These results too come from a model that includes all covariates listed in Table 1. The full model can be found in Appendix 6 as well as the MCMC diagnostics of the significant parameter (see Table 6.1 and Figure 6.1).

The transformed coefficient for the significant variable in this model (driver education mandatory in the learner stage) is 0.655 (REML p-value=0.036) and can be interpreted as follows. The relative fatality risk of 18 year old drivers in jurisdictions where driver education is mandatory in the learner phase decreases by a factor of 0.655, or by 34.5% ((1-0.655)*100%), compared to those jurisdictions where driver education is not mandatory in the learner phase. This variable is significant according to both estimation procedures.

The model for 18 year old drivers was further investigated by checking for outliers and the relative influence of individual jurisdictions. No single jurisdiction had an effect that was particularly greater than that of the other jurisdictions and no outliers were found (see Figure 6.2 in Appendix 6).

This significant covariate, however, does not explain any of the variance between jurisdictions (i.e., heterogeneity) according to the adjusted R-squared statistic of this model and 11.43% of the residual variation is due to heterogeneity (see Table 6.1). This extreme value for R-squared can be explained by the fact that the model may not satisfactorily fit the data. This can be derived from the value of tau-squared, which is equal to zero. As such, this R-squared statistic bears no meaning and it is recommended not to rely on it. Note that the coefficient of the significant effect in this model is not necessarily adversely affected by this lack of model fit. In this regard, it is interesting to see that the Bayesian estimates — that are not affected at all by the potential lack of fit of the REML model — are very comparable to the REML estimates.

Finally, the Bayesian DIC was 110.13 for the random model while it was 115.82 for the fixed model. The lower value for the random model confirms such a model is indeed more appropriate than a fixed model because it has a better fit.

Meta-regression for 19 year old drivers

Table 2 provides an overview of the significant effects that were found for 19 year old drivers and compares the results from both estimation procedures. These results too come from a model that includes all covariates listed in Table 1. The full model can be found in Appendix 7, as well as the MCMC diagnostics per significant parameter (see Table 7.1 and Figure 7.1).

When checking this model for outliers and the relative influence of individual jurisdictions, one outlier was identified (Maryland) — see Figure 7.2 in Appendix 7. Therefore, the results for 19 year old drivers are based on a model that excludes this outlier.

As can be seen, five effects were significant according to the MCMC Gibbs estimation procedure, but only two effects were significant according to the REML results. The first variable is length of night restriction in the learner stage. This variable is significant according to both estimation procedures. Its transformed coefficient is 1.11 (REML p-value=0.038). This means that for an increase in length of the night restriction in the learner stage of one hour, the relative fatality risk of 19 year old drivers increases by 11% ((1.11-1)*100).

The second variable is country and its transformed coefficient is 13.29. As can be seen in Table 2, this variable is only significant according to the MCMC Gibbs results (REML p-value=0.074). Its interpretation is as follows. The relative fatality risk of 19 year old drivers in Canadian jurisdictions is 1,229% ((13.29-1)*100) higher than that of 19 year old drivers in U.S. jurisdictions. Apparently, there is something different between Canada and the U.S. that has negative consequences for the relative fatality risk of 19 year old drivers in Canadian jurisdictions.

The transformed coefficient for the third variable is 52.09. This variable is also not significant according to the REML estimates (REML p-value=0.072). It means that the relative fatality risk of 19 year old drivers in jurisdictions where night restrictions are lifted for work purposes in the intermediate stage increases by a factor of 52.09 or 5,109% ((52.09-1)*100%), compared to the relative fatality risk of 19 year old drivers in those jurisdictions where such night restrictions are not lifted.

The fourth variable (driver education) is only significant according to the MCMC Gibbs estimates. Its transformed coefficient is 2.11 (REML p-value=0.073). According to this variable there is an increase in the relative fatality risk of 19 year old drivers of 111% in jurisdictions with mandatory driver education in the intermediate stage ((2.11-1)*100).

The transformed result for the last variable (exit test) is 0.02 (REML p-value=0.026). This variable is significant according to both estimation procedures. The interpretation of this coefficient implies that the relative fatality risk of 19 year old drivers in jurisdictions that require an exit test to graduate from the intermediate stage is 0.02 times smaller, or 98% ((1-0.02)*100) than that of 19 year old drivers in jurisdictions that do not require such an exit test.

According to the adjusted R-squared statistic of this model all the variance between jurisdictions or heterogeneity is explained (which implies that none of the residual variation is due to heterogeneity) — see Table 7.1. This extreme value for R-squared can again be explained by the fact that the model does not fit the data well enough for the R-squared statistic to be meaningful. The tau-squared statistic is also equal to zero. As a consequence, it is not recommended to rely on these model fit statistics. Note again that the coefficients of the significant effects in this model are not necessarily adversely affected by this lack of model fit. Interestingly, the Bayesian estimates are again very comparable to the REML estimates.

Finally, the Bayesian DIC was 71.85 for the random model while it was 76.19 for the fixed model. The lower value for the random model confirms such a model is indeed more appropriate than a fixed model because it has a better fit.

DISCUSSION

Limitations

When interpreting the results from the analyses in this study, some limitations of the applied methods have to be borne in mind. First, the scope of this study was limited geographically because only data from North-American jurisdictions were used, while other GDL programs exist outside of North America. However, the decision to include raw data from all North-American jurisdictions, rather than to use only published results of the relatively few GDL programs that have been evaluated, considerably broadened the scope of this study. Furthermore, if a more restricted meta-analysis had been carried out based on a systematic review of available evaluation studies, the study's scope would probably also have been limited to North America because the bulk of the available literature comes from evaluations in those jurisdictions.

Second, all the analyses are age-based. As such, it is implicitly assumed that 16, 17, 18 and 19 year old drivers are affected by GDL while 25-54 year old drivers are not. While such an assumption is true to a large extent, some bias may have been introduced in the analyses due to adopting such an age-based approach. For example, in most Canadian and a few U.S. jurisdictions GDL applies to all novice drivers, regardless of their age, which means that for those jurisdictions the comparison group of 25-54 year old drivers may contain some drivers that have been affected by GDL. Also, the denominator used in this study to calculate the relative fatality risks is the entire population of a particular age, rather than the number of licensed drivers of that age. It can be argued that the population truly at risk is the number of licensed drivers rather than the entire population. However, it is extremely difficult, if not impossible to obtain such information about the number of licensed drivers.

The post period in this study began 12 months after implementation of a GDL program until 24 months after implementation. As such, it can be expected that young drivers aged 16 and 17 would have gone through the new GDL system, as well as many 18 year old drivers and some 19 year old drivers. In Canada and a few U.S. states, GDL applies to new drivers of all ages and not all teenagers apply for a license immediately after they become eligible but may wait until they turn 17, 18 or 19. This means the evaluation period that was chosen likely captured young drivers of any age (i.e., 16, 17, 18 and 19) who have been exposed to the new GDL system. However, it is acknowledged that such an evaluation period may perhaps not have been long enough for a sufficient number of 19 year old drivers (and perhaps 18 year old drivers too) to already have gone through this new GDL system. This may explain why results showed no summary effect of GDL on these older teens.

Furthermore, this age-based approach does not allow disentangling the effects of the different GDL stages to the same extent that an analysis based on license status would. This also makes it difficult sometimes to explain the reasons why an effect emerges. For example, an increase in crashes after lifting passenger limits when passengers are immediate family members could be due to the negative influence of too many passengers on inexperienced drivers, but it could also be due to exposure, i.e., perhaps novice drivers simply drive more often when such a passenger limit is lifted, which would expose them more to opportunities for crashing. The same is true for the effect that was found with respect to lifting night restrictions for work purposes. While ambitious, challenging, and perhaps not feasible, the results obtained in this study may improve and could be further explained if comparable analyses could be conducted using license status-based data, rather than age-based data, or if these results could be related to exposure data. On the other hand, some would argue that such age-based per capita rates are actually better suited to capture the overall effects of GDL.

Third, only fatality data were used to investigate the impact of GDL while GDL may perhaps have a more profound effect on injury and property damage only (PDO) crashes involving young drivers because injury and PDO crashes are so much more prevalent than fatal crashes. The decision to use fatality data, however, was made in light of the search for a common denominator among all North-American jurisdictions and because only standardized fatality data for the U.S. were available.

Fourth, as explained in the method section, certain jurisdictions were included in the analyses more than once to reflect any legislative changes that were passed after the implementation of a GDL program. As such, the final database contained information from 78 ‘jurisdictions’, rather than from 58. The upside of such an approach is that more data are used, which improves precision of the estimates and also enables a more nuanced investigation of differences between programs — when a program is improved it should logically lead to an additional reduction in fatalities, above and beyond the already established reduction due to the implementation of the original program, assuming other relevant circumstances in the jurisdiction have not changed too much. However, the downside is that the observations in the database (or ‘data points’) for such jurisdictions are dependent inasmuch as improvements to a program are conditional on the features of the already established program. While this may have adversely affected the results, it should be noted that all the data points in the database are dependent, at least to some extent, because it is unlikely that a GDL program would have been implemented in a vacuum, entirely independently from what is being done in other jurisdictions.

In this regard, it had previously been suggested that there may be a learning effect in that jurisdictions that are moving forward with the implementation of a GDL program can benefit from lessons learned in other, pioneering jurisdictions; or, that earlier GDL programs can create a more receptive climate for the implementation of stronger GDL programs. To further bolster this possibility, the different GDL programs were sorted chronologically and for each program the number of components that can make a program more effective were counted. The results can be found in Appendix 8, Table 8.1. As can be seen, the number of components increases gradually over time. For example, the average number of components until 1999, inclusive is 5.9, while the average number from 2000 until 2005, inclusive is 7.8. While a significant learning effect was found in earlier versions of the model with 16 year old drivers during an exploratory data analysis phase, unfortunately this could not be confirmed with any of the final models. As such, when looking at how programs evolved over time there seems to be some evidence to suggest there is a need for dissemination of study results as other jurisdictions may benefit from it, but it was not possible to confirm this formally with the meta-analyses.

Finally, while information from 78 jurisdictions was collected, the true sample size of the meta-regression analyses was only 48, due to missing data for several of the covariates. It proved extremely challenging and very labor-intensive to obtain complete GDL data records for each of the jurisdictions included in this study, despite the availability of information about these programs on two on-line inventories (IIHS's and TIRF's) and despite the access to experts in the field.

Findings

Using summary effects coming from a random effects DerSimonian and Laird model, strong evidence in support of GDL reducing fatalities was found. The evidence, however, only showed that GDL has had a positive and significant impact on fatality rates among 16 year old drivers, when adjusting for a group of older drivers who are assumed not to have been affected by the implementation of GDL (reduction in the relative fatality risk of 19.1%). No evidence was found to suggest GDL has had an overall impact on relative fatality risks among 17, 18 and 19 year old drivers when looking at the summary effects only.

The results also showed that there was some heterogeneity among jurisdictions that could provide insight into how GDL has significantly decreased fatalities among 16, 18 and 19 year old drivers. None of the variance in the outcome measure for 17 year old drivers, however, was due to heterogeneity, which is hard to explain. Compared to the patterns of variance that emerged among 16, 18 and 19 year old drivers (5.1%, 11.1% and 5.5% respectively), this lack of variance among 17 year old drivers seems like an anomaly, more than anything else. If it would be assumed that there truly is no variance among 17 year old drivers, this would mean that all GDL programs are equally effective for this age group. Given that most GDL programs differ considerably from one another and that 17 year old drivers in this study may have been in the learner, intermediate, or full license stages due to the nature of population data, it seems highly unlikely that this assumption holds true. Perhaps the composition of this group differs from the other age groups in that there is a greater variety of drivers in terms of license status among 17 years old drivers than there is among 16, 18 and 19 year old drivers. For example, the majority of 16 year old drivers may be in the learner stage, and the majority of 18 and 19 year old drivers may be in the full stage, whereas all stages, including the intermediate stage, could be equally represented among 17 year old drivers. However, such an explanation cannot be tested due to the limitations of age-based data and, as a consequence, is only speculative.

As for 18 and 19 year old drivers, the summary effect may not have suggested GDL has had a positive impact, but the proportion of variance due to heterogeneity (11.1% and 5.5% respectively) suggested it may still be possible to distinguish GDL programs which have had a significant and positive impact on relative fatality risks among 18 and 19 year old drivers from those GDL programs that have not had such an impact.

Two variables were significant for 16 year old drivers, namely passenger restrictions in the intermediate stage and whether passenger restrictions are lifted or not in the intermediate stage if passengers are family members. Both variables were significant according to both estimation procedures (REML and MCMC Gibbs).

The first variable suggests that passenger restrictions in the intermediate stage are beneficial in that such restrictions lead to a 88.5% decrease in the relative fatality risk of 16 year old drivers. Conversely, the second variable suggests it is beneficial not to lift passenger limits in the intermediate stage if passengers are family members because lifting such restrictions leads to an increase in the relative fatality risk of 728.1%. Even though it could be argued that some of the 16 year old drivers are in the learner stage, certainly there are 16 year old drivers who have already graduated to the intermediate stage. This finding may suggest that restrictions should not only be license status based but instead be based on a combination of license status and age as lifting such a restriction for 16 year olds, even if they have already graduated to a more advanced stage, proves to have negative consequences.

One variable (mandatory driver education in the learner stage) was significant with 18 year old drivers according to both estimation procedures. Mandatory driver education in the learner stage may be beneficial in that it can lead to a 34.5% reduction in the relative fatality risk of 18 year old drivers. It is noteworthy that this finding stands in marked contrast to what has previously been reported in the literature (see for example Mayhew 2007 for an overview). Keeping the limitations of this study in mind, such a finding should be further investigated.

Five variables were significant for 19 year old drivers, but not all of them were significant according to both estimation procedures. The first variable (length of night restriction in the learner stage) was significant according to both procedures. According to the results for this variable, an increase in the length of the night restriction in the learner stage of one hour leads to an 11% increase of the relative fatality risk of 19 year old drivers. Assuming the majority of drivers in this age cohort would be in the intermediate or full stage, this may mean that too little night-time driving practice in the learner stage may have negative consequences for drivers entering the less restricted intermediate or full stage, although the limitations of aged based data, as discussed in the previous section, do not allow confirming whether such a residual effect is truly at work or not. For this to be true, a sufficient number of 19 year old drivers would have had to have gone through the new GDL licensing system, which is not certain given the post period that was used in this study. The status of this interpretation is therefore only speculative.

The variable country was only significant according to one procedure. According to this variable, the relative fatality risk of 19 year old drivers in Canadian jurisdictions is 1,229% higher than that of 19 year old drivers in U.S. jurisdictions. Because the results for this variable are controlled for numerous other variables that describe GDL programs, it is likely that other non-GDL related features, specific to Canadian jurisdictions and different from U.S. jurisdictions may explain this extremely high increase in the relative fatality risk of 19 year old drivers in Canadian jurisdictions. Further investigation is needed to better understand this result.

Evidence was found against lifting the night restrictions in the intermediate stage for work purposes, as it is associated with a 5,109% increase in the relative fatality risk of 19 year old drivers. This variable, however, was only significant according to one procedure. It seems GDL programs could be enhanced if stricter nighttime driving restrictions were applied. This is consistent with previous findings (see Williams 2007).

The fourth variable (mandatory driver education in the intermediate stage) also was only significant according to one estimation procedure. Its interpretation is as follows. There is an increase in the relative fatality risk of 19 year old drivers of 111% in jurisdictions with mandatory driver education in the intermediate stage. While this finding may seem counterintuitive, it is more in line with other research (see Mayhew 2007). Further investigation is needed to better understand this result.

Finally, requiring an exit test to graduate from the intermediate stage is beneficial since it leads to a 98% reduction in the relative fatality risk.

It is noteworthy that model fit for each model was further investigated using a variety of model fit statistics, including thvve proportion of residual variation due to heterogeneity, the proportion of heterogeneity explained by the covariates in the model (adjusted R-squared), the Bayesian DIC and checks for outliers and the relative influence of individual jurisdictions. Model fit statistics for the model with 16 year old drivers indicated a very good fit, while model fit was less good for 18 and 19 year old drivers. Taken together, however, these statistics suggest that the models for 18 and 19 year old drivers fit the data reasonably well. In this regard, it was interesting to see convergence between the coefficients coming from both estimation procedures (REML and MCMC Gibbs).

CONCLUSION AND RECOMMENDATIONS

In conclusion, the meta-analysis used in this study has proven to be a useful approach. Despite some limitations of the study design, some previously established findings have been confirmed and some interesting and intriguing new findings emerged from these analyses. Some of the results were only significant according to one approach and perhaps they should only serve as exploratory findings that need further investigation and confirmation. Other results were strong and highly significant according to both approaches that were adopted. Such results should also be further investigated and confirmed.

Some caution is warranted when interpreting the results from this study. For example, some effects appeared to be very strong, but the accompanying confidence intervals were wide. As such, real effects probably exist, but more research is needed to more reliably estimate the strength of these effects.

Further research along these lines seems promising. Based on this study, the following recommendations are formulated:

- It should be investigated in a prospective study if it would be feasible to conduct a large scale project to replicate the findings of the present study using license status based data rather than age-based data and also to estimate at what cost this could be done. This would imply gaining access to driver records systems of each of the jurisdictions included in this study (and preferably jurisdictions elsewhere in the world) and would involve sophisticated and complex data manipulation. This seems very ambitious and such an approach could probably only succeed by relying on several teams of data-analysts who would collaborate under the guidance of a coordinator. If feasible, however, such a project could produce highly relevant and important results and overcome several of the limitations of aged-based data described previously.
- In the interim, a follow-up study should be conducted to complete data records of several jurisdictions and to refine the formal description of the GDL programs by means of covariates. Collecting more data for each of the covariates with missing values could increase precision of the estimates and reveal more patterns. Refining the formal description of GDL programs could also reveal more patterns between independent and dependent variables. This could be accomplished by establishing a panel of experts who could discuss and refine the definitions of the covariates, elaborate on the list of covariates, and use their resources to obtain more data. Comparable analyses could be re-run and could lead to uncovering more patterns, useful for the improvement of GDL programs.
- Other avenues for further research include replicating the models by gender; using Canadian jurisdictions to model effectiveness of GDL programs in terms of fatalities and injuries instead of just fatalities; using longer evaluation periods (e.g., 24 months rather than 12 months); and, re-analyzing the data excluding second and further data points for the same jurisdiction.

REFERENCES

Altman, D.G., Bland, J.M. 2003. Interaction revisited: the difference between two estimates. *British Medical Journal* 326, 219.

Baker, S.P., Chen, L.H., Li, G., 2006. National Evaluation of Graduated Driver Licensing Programs. NHTSA, Washington, DC.

Baker, S.P., Chen, L.H., and Li, G. 2007. Nationwide Review of Graduated Driver Licensing. Washington, DC: AAA Foundation for Traffic Safety.

Browne, W.J., 2004. MCMC estimation in MLwiN. Version 2.0. Updated version for 2.0 by William J. Browne and Edmond S.W. Ng. Centre for Multilevel Modeling. Institute of Education. London: University of London.

Carlin, B.P., Louis, T.A. 2000. Bayes and Empirical Bayes Methods for Data Analysis. Washington, DC: Chapman & Hall/CRC.

Clarke, M.J., Stewart, L.A. 2007. Obtaining individual patient data from randomised controlled trials. In: M. Egger, G.D. Smith and D.G. Altman (eds.). Systematic Reviews in Health Care: Meta-Analysis in Context (second edition). pp. 109-121. London: BMJ Publishing Group.

Dee, T.S., Grabowski, D.C., Morrisey, M.A., 2005. Graduated driver licensing and teen traffic fatalities. *Journal of Health Economics* 24, 571-589.

Deeks, J.J., Altman, D.G., Bradburn, M.J. 2007. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: M. Egger, G.D. Smith and D.G. Altman (eds.). Systematic Reviews in Health Care: Meta-Analysis in Context (second edition). pp. 285-312. London: BMJ Publishing Group.

Egger, M., Smith, G.D. 2007. Principles of and procedures for systematic reviews. In: M. Egger, G.D. Smith and D.G. Altman (eds.). Systematic Reviews in Health Care: Meta-Analysis in Context (second edition). pp. 23-42. London: BMJ Publishing Group.

Elvik, R., 2005a. Can We Trust the Results of Meta-Analyses? A Systematic Approach to Sensitivity Analysis in Meta-Analysis. *Journal of the Transportation Research Board*, No. 1908, 221-229.

Elvik, R., 2005b. Introductory Guide to Systematic Reviews and Meta-Analysis. *Journal of the Transportation Research Board*, No. 1908, 230-235.

Foss, R.D., Evenson, K.R., 1999. Effectiveness of Graduated Driver Licensing in Reducing Motor Vehicle Crashes. *American Journal of Preventive Medecine* 16 (1S), 47-55.

Foss, R., Feaganes, J.R., Rodgman, E.A. 2001. Initial effects of Graduated Driver Licensing on 16-Year-Old driver crashes in North Carolina. Journal of the American Medical Association 286 (13), 1588-1592.

Goldstein, H., 2003. Multilevel statistical models. London: Arnold.

Harbord, R.M., Higgins, J.P.T. (2008). Meta-regression in Stata. *The Stata Journal*, 8(4), 493-519.

Hartling, L., Wiebe, N., Russell, K., Petruk, J., Spinola, C., Klassen, T.P., 2005. Graduated driver licensing for reducing motor vehicle crashes among young drivers. The Cochrane Database of Systematic Reviews, Issue 2. Art. No.: CD003300.pub2.

Hedlund, J., Shults, R.A., Compton, R., 2003. What we know, what we don't know, and what we need to know about graduated driver licensing. *Journal of Safety Research* 34, 107-115.

Hedlund, J., Shults, R.A., Compton, R., 2006. Graduated driver licensing and teenage driver research in 2006. *Journal of Safety Research* 37, 107-121.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G. 2003. Measuring inconsistency in meta-analysis. *British Medical Journal* 327, 557-560.

Hox, J.J., de Leeuw, E.D., 2003. Multilevel models for Meta-Analysis. In: S.P. Reise & N. Duan (Eds.). *Multilevel modeling. Methodological Advances, Issues, and Applications.* Mahwah, N.J.: Lawrence Erlbaum Associates.

Lambert, P.C., Abrams, K.R. 1995. Meta-Analysis Using Multilevel Models. *Multilevel Modelling Newsletter* 7(2), 16-18.

Mayhew, D.R. (2007). Driver education and graduated licensing in North America: Past, present and future. Journal of Safety Research 38, 229-235.

Mayhew, D.R., Simpson, H.M., Groseillers M. 1999. Impact of the graduated driver licensing program in Nova Scotia. Ottawa: Traffic Injury Research Foundation.

Mayhew, D.R., Simpson, H.M., Groseillers M., Williams, A.F. 2001. Impact of the graduated driver licensing program in Nova Scotia. *Journal of Crash Prevention and Injury Control* 2(3), 179-192.

Mayhew, D.R., Simpson, H.M., Williams, A.F., and Desmond, K. 2002. Specific and Long-Term Effects of Nova Scotia's Graduated Licensing Program. Arlington, VA: Insurance Institute for Highway Safety.

Mayhew, D.R., Simpson, H.M, Singhal, D., 2005. Best Practices for Graduated Driver Licensing in Canada. Ottawa, Ontario: Traffic Injury Research Foundation.

Morrisey, M.A., Grabowski, D.C., Dee, T.S., Campbell, C., 2006. The strength of graduated drivers license programs and fatalities among teen drivers and passengers. *Accident Analysis and Prevention* 38 (1), 135-141.

Raftery, A.E., Lewis, S.M. 1992. How many iterations in the Gibbs Sampler? In: J.M. Bernardo, J.O. Berger, A.P. David and A.F.M. Smith (eds.). *Bayesian Statistics* 4, pp. 763-773. Oxford: Oxford University Press.

Shope, J.T., Molnar, L.J., Elliott, M.R., Waller, P.F. 2001. Graduated Driver Licensing in Michigan: Early impact on motor vehicle crashes among 16-year-old drivers. *Journal of the American Medical Association* 286 (13), 1599-1598.

Shope, J.T. 2007. Graduated driver licensing: Review of Evaluation results since 2002. *Journal of Safety Research* 38(2), 165-175.

Simpson, H.M., 2003. The evolution and effectiveness of graduated licensing. *Journal of Safety Research* 34, 25-34.

Smith, T.C., Spiegelhalter, D.J., Thomas, A. 1995. Bayesian Approaches to Random-Effects Meta-Analysis: A Comparative Study. *Statistics in Medicine* 14, 2685-2699.

StataCorp. (2007) Stata Statistical Software. Release 10. College Station, Texas: StataCorp LP.

Statistics Canada 2007. Demographics Estimates Compendium 2007 (CD-ROM). 91-213-SCB.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B* 64, 583-640.

Sterne, J.A.C., Bradburn, M.J., Egger, M. 2007a. Meta-analysis in Stata. In: M. Egger, G.D. Smith and D.G. Altman (eds.). Systematic Reviews in Health Care: Meta-Analysis in Context (second edition). pp. 347-372. London: BMJ Publishing Group.

Sterne, J.A.C, Egger, M., Smith, G.D. 2007b. Investigating and dealing with publication and other biases. In: M. Egger, G.D. Smith and D.G. Altman (eds.). Systematic Reviews in Health Care: Meta-Analysis in Context (second edition). pp. 347-372. London: BMJ Publishing Group.

Thompson, S.G. 2007. Why and how sources of heterogeneity should be investigated. In: M. Egger, G.D. Smith and D.G. Altman (eds.). Systematic Reviews in Health Care: Meta-Analysis in Context (second edition). pp. 157-175. London: BMJ Publishing Group.

Turner, R.M., Omar, R.Z., Yang, M., Goldstein, H., Thompson, S.G. 1999. Random effects meta-analysis of trials with binary outcomes using multilevel models in MLwiN. *Multilevel Modelling Newsletter* 11(1), 6-9.

Ulmer, R.G., Preusser, D.F., Williams, A.F., Ferguson, S.A., Farmer, C.M. 2000. Effect of Florida's graduated licensing program on the crash rate of teenage drivers. *Accident Analysis and Prevention* 32(4), 527-532.

Williams, A.F., 2003. The compelling case for graduated licensing. *Journal of Safety Research* 34, 3-4.

Williams, A.F., 2007. Contribution of the components of graduated licensing to crash reductions. *Journal of Safety Research* 38, 177-184.

Zwicker, T.J., Williams, A.F., Chaudhary, N.K., Farmer, C.M., 2006. Evaluation of California's Graduated Licensing System. Arlington, VA: Insurance Institute for Highway Safety.

APPENDIX

Appendix 1: Summary effect for 16 year old drivers

Table 1.1: Random Effects Meta-Analysis

jurisdiction	ES	[95% Conf.	Interval]	% Weight
AK	1.914	0.061	60.256	0.13
AL	1.272	0.627	2.581	2.79
AR1	0.852	0.311	2.338	1.45
AR2	0.815	0.290	2.293	1.39
CA	1.351	0.738	2.473	3.67
CO	0.224	0.048	1.052	0.64
CT1	0.603	0.098	3.699	0.47
CT2	1.185	0.073	19.342	0.20
DC	0.275	0.005	14.926	0.10
DE	0.914	0.142	5.867	0.45
FL	3.389	1.459	7.873	2.03
GA1	0.629	0.344	1.148	3.70
GA2	1.283	0.722	2.278	4.01
IA	0.574	0.190	1.739	1.22
ID	0.999	0.061	16.296	0.20
IL	1.000	0.521	1.922	3.22
IN	0.668	0.344	1.298	3.12
KY1	0.357	0.154	0.832	2.02
LA	0.889	0.362	2.183	1.80
MA1	1.341	0.412	4.359	1.08
MA2	2.103	0.511	8.651	0.76
MD1	1.086	0.216	5.464	0.59
MD2	0.544	0.160	1.849	1.01
ME1	1.186	0.159	8.852	0.38
ME2	6.405	0.329	124.496	0.18
ME3	0.273	0.031	2.435	0.32
MI	0.479	0.233	0.982	2.71
MN1	1.901	0.580	6.230	1.07
MN2	1.142	0.461	2.831	1.77
MO	0.682	0.349	1.331	3.08
MS	1.239	0.556	2.765	2.22
NC	0.258	0.110	0.604	2.00
ND	1.757	0.056	54.994	0.13
NE	1.458	0.405	5.253	0.92
NH1	1.049	0.162	6.780	0.44
NH2	2.504	0.244	25.693	0.29
NH3	1.075	0.224	5.145	0.62
NJ	0.460	0.015	13.841	0.13
NM	0.243	0.011	5.452	0.16
NV	1.013	0.220	4.660	0.66
NY1	1.169	0.435	3.147	1.51
NY2	0.541	0.183	1.603	1.26
OH1	0.486	0.226	1.046	2.42
OH2	0.626	0.276	1.422	2.14
OK	1.899	0.733	4.924	1.62
OR	0.259	0.030	2.252	0.33
PA	0.515	0.237	1.115	2.38
RI1	1.257	0.073	21.539	0.19
RI2	2.002	0.064	62.588	0.13
SC1	0.722	0.349	1.493	2.66
SC2	0.472	0.159	1.397	1.26
SD	0.545	0.018	16.781	0.13
TN	0.546	0.284	1.051	3.21
TX	0.640	0.393	1.045	5.22
UT1	2.297	0.430	12.262	0.55

```
UT2                  |  2.061       0.333    12.750           0.46
UT3                  |  0.509       0.125     2.063           0.78
VA1                  |  0.703       0.320     1.543           2.31
VA2                  |  1.594       0.631     4.027           1.70
VA3                  |  0.827       0.345     1.985           1.90
VA4                  |  0.836       0.379     1.847           2.28
VT                   |  0.266       0.028     2.543           0.30
WA                   |  3.318       0.679    16.221           0.61
WI                   |  0.684       0.295     1.589           2.03
WV                   |  0.630       0.164     2.416           0.84
WY                   |  0.333       0.010    10.615           0.13
Alberta              |  0.203       0.057     0.716           0.95
BC1                  |  1.504       0.594     3.811           1.69
BC2                  |  0.354       0.071     1.772           0.59
Manitoba             |  0.200       0.010     3.896           0.18
NewBrunswick         |  0.475       0.016    14.553           0.13
Newfoundland         |  1.174       0.036    37.906           0.13
NovaScotia           |  0.458       0.046     4.555           0.29
Ontario              |  0.598       0.299     1.197           2.89
PrinceEdwardI        |  0.430       0.033     5.691           0.23
Quebec               |  1.329       0.443     3.984           1.24
Saskatchewan         |  1.811       0.057    57.181           0.13
Yukon                |  2.900       0.042   199.675           0.09
---------------------+---------------------------------------------------
D+L pooled ES        |  0.809       0.714     0.917         100.00
---------------------+---------------------------------------------------

  Heterogeneity chi-squared =  81.17 (d.f. = 77) p = 0.351
  I-squared (variation in ES attributable to heterogeneity) =   5.1%
  Estimate of between-study variance Tau-squared =  0.0158

  Test of ES=1 : z=   3.32 p = 0.001
```

Figure 1.1: Forest Plot

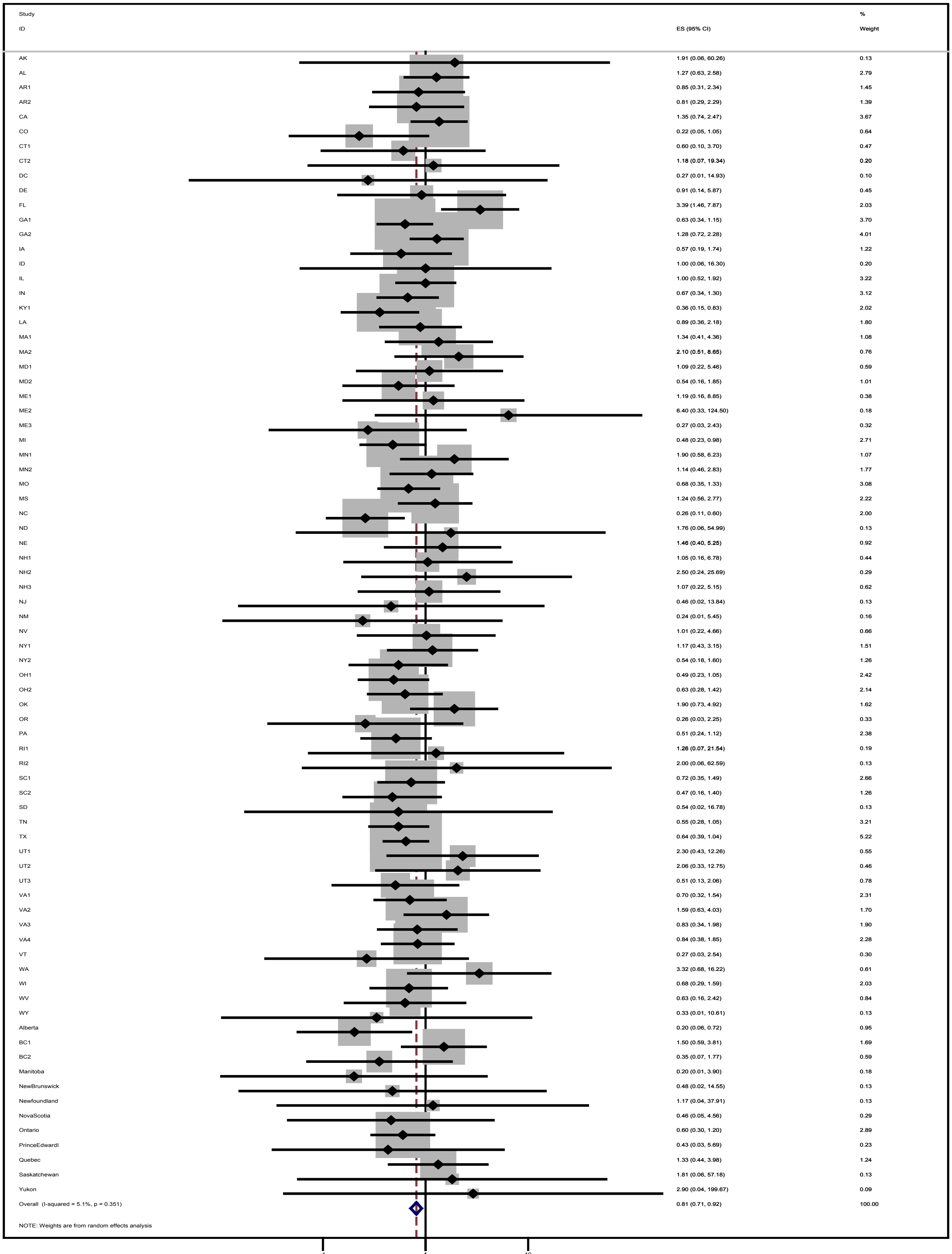
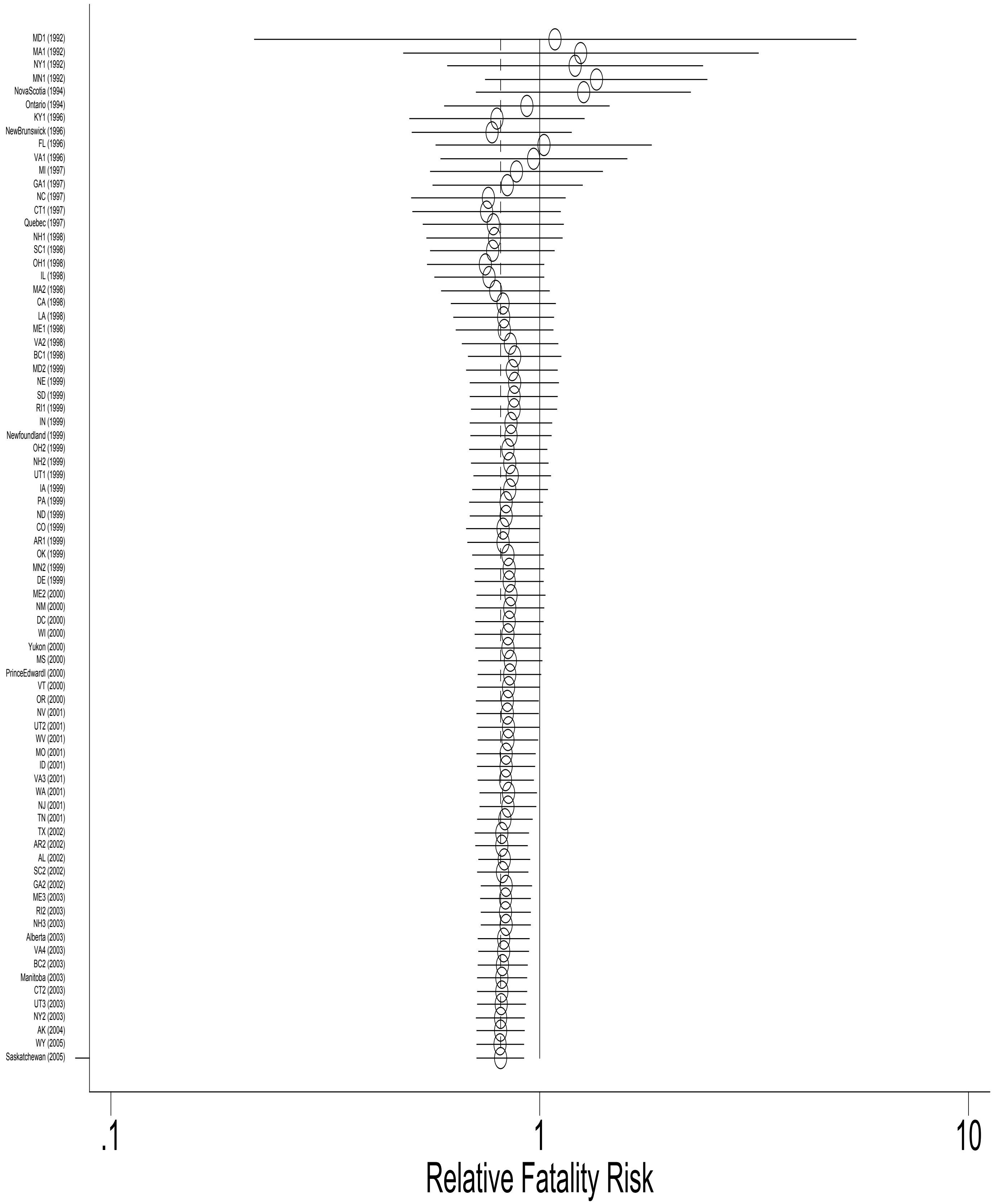


Figure 1.2: Cumulative Random Effects Meta-Analysis



Appendix 2: Summary effect for 17 year old drivers

Table 2.1: Random Effects Meta-Analysis

Jurisdiction	ES	[95% Conf. Interval]		% Weight
AK	0.245	0.010	5.764	0.10
AL	1.223	0.569	2.626	1.69
AR1	0.852	0.376	1.927	1.48
AR2	1.172	0.519	2.648	1.49
CA	1.132	0.677	1.892	3.75
CO	0.874	0.310	2.464	0.92
CT1	0.708	0.154	3.257	0.42
CT2	1.595	0.343	7.408	0.42
DC	0.383	0.007	20.828	0.06
DE	0.712	0.121	4.185	0.32
FL	1.308	0.751	2.280	3.20
GA1	0.795	0.454	1.393	3.14
GA2	1.465	0.806	2.663	2.77
IA	0.787	0.288	2.148	0.98
ID	0.984	0.236	4.099	0.49
IL	1.089	0.584	2.034	2.54
IN	1.031	0.519	2.048	2.10
KY1	1.148	0.589	2.238	2.22
LA	0.979	0.391	2.452	1.17
MA1	0.498	0.146	1.706	0.65
MA2	1.464	0.564	3.800	1.09
MD1	0.250	0.053	1.177	0.41
MD2	1.652	0.693	3.942	1.31
ME1	1.828	0.289	11.552	0.29
ME2	0.589	0.175	1.988	0.67
ME3	0.694	0.163	2.957	0.47
MI	1.649	0.876	3.104	2.47
MN1	0.365	0.152	0.877	1.28
MN2	1.021	0.444	2.347	1.43
MO	0.672	0.369	1.224	2.76
MS	1.217	0.597	2.484	1.94
NC	0.891	0.488	1.627	2.73
ND	0.423	0.112	1.605	0.56
NE	0.618	0.154	2.482	0.51
NH1	0.181	0.008	4.194	0.10
NH2	0.253	0.025	2.591	0.18
NH3	1.072	0.224	5.135	0.40
NJ	2.843	0.753	10.742	0.56
NM	2.520	0.478	13.279	0.36
NV	0.972	0.252	3.741	0.54
NY1	1.017	0.525	1.971	2.26
NY2	0.855	0.473	1.547	2.81
OH1	1.183	0.569	2.459	1.85
OH2	0.841	0.458	1.546	2.67
OK	1.059	0.435	2.575	1.25
OR	1.037	0.271	3.969	0.55
PA	0.718	0.390	1.321	2.66
RI1	0.582	0.018	18.344	0.08
RI2	0.687	0.105	4.495	0.28
SC1	0.867	0.416	1.807	1.83
SC2	1.471	0.702	3.086	1.80
SD	0.222	0.025	2.001	0.20
TN	0.933	0.484	1.797	2.30
TX	1.025	0.682	1.541	5.94
UT1	0.301	0.059	1.544	0.37
UT2	0.981	0.296	3.252	0.69

```
UT3                  |  0.202       0.024     1.733          0.21
VA1                  |  0.677       0.308     1.487          1.60
VA2                  |  1.457       0.738     2.875          2.14
VA3                  |  0.733       0.364     1.477          2.02
VA4                  |  1.345       0.605     2.990          1.55
VT                   |  2.231       0.072    69.485          0.08
WA                   |  1.298       0.508     3.312          1.13
WI                   |  0.765       0.355     1.648          1.68
WV                   |  0.691       0.204     2.338          0.66
WY                   |  0.707       0.013    37.759          0.06
Alberta              |  0.928       0.361     2.388          1.11
BC1                  |  1.184       0.544     2.577          1.64
BC2                  |  1.314       0.534     3.236          1.22
Manitoba             |  1.250       0.316     4.950          0.52
NewBrunswick         |  1.897       0.166    21.714          0.17
Newfoundland         |  0.098       0.004     2.138          0.10
NovaScotia           |  1.940       0.413     9.122          0.41
Ontario              |  1.113       0.658     1.881          3.58
PrinceEdwardI        |  0.909       0.016    51.292          0.06
Quebec               |  1.592       0.835     3.036          2.37
Saskatchewan         |  0.235       0.010     5.544          0.10
Yukon                |  5.574       0.131   237.211          0.07
---------------------+----------------------------------------------------
D+L pooled ES        |  1.001       0.906     1.105        100.00
---------------------+----------------------------------------------------

  Heterogeneity chi-squared =  53.16 (d.f. = 77) p = 0.983
  I-squared (variation in ES attributable to heterogeneity) =   0.0%
  Estimate of between-study variance Tau-squared =  0.0000

  Test of ES=1 : z=   0.02 p = 0.988
```

Figure 2.1: Forest Plot

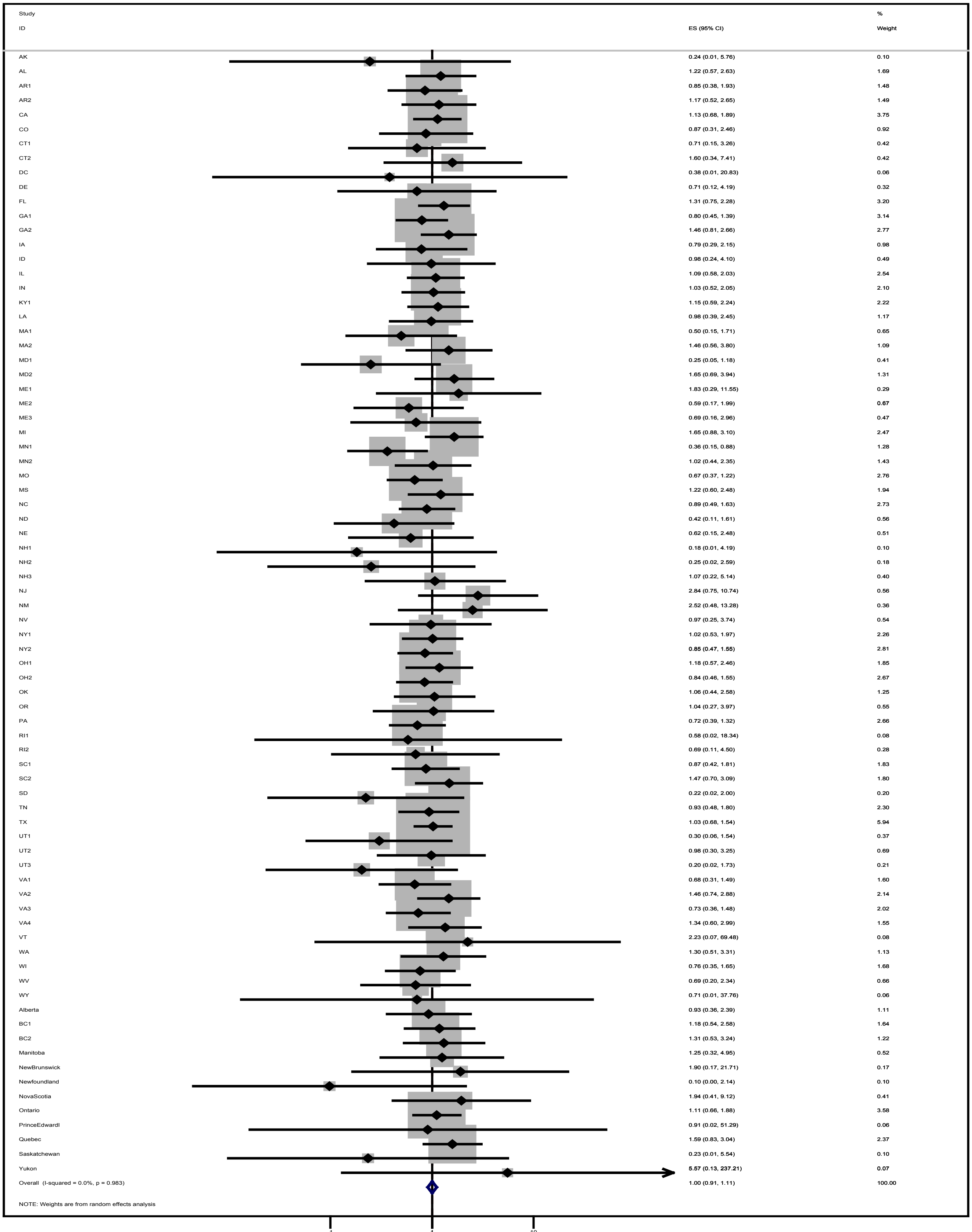
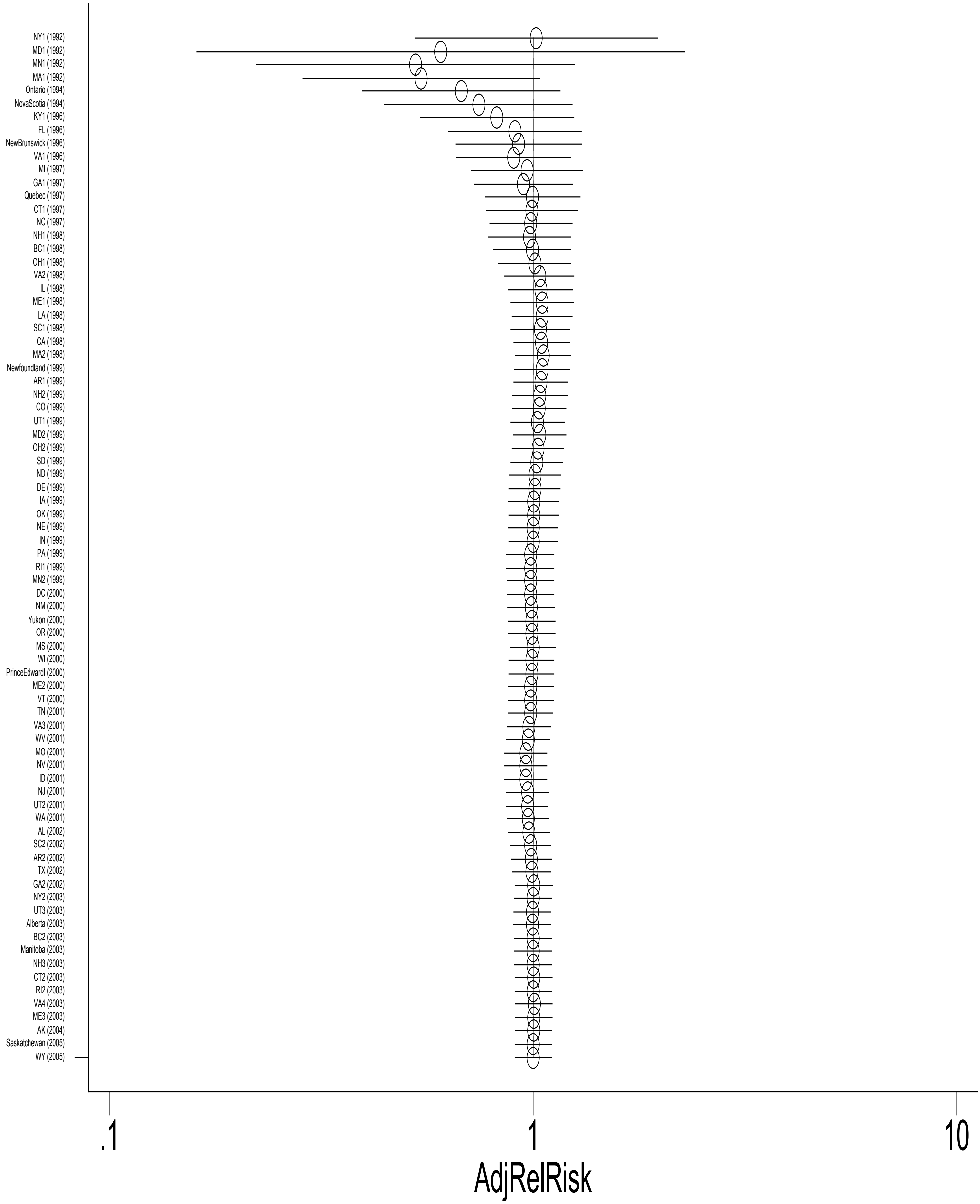


Figure 2.2: Cumulative Random Effects Meta-Analysis



Appendix 3: Summary effect for 18 year old drivers

Table 3.1: Random Effects Meta-Analysis

Jurisdiction	ES	[95% Conf. Interval]		% Weight
AK	0.990	0.019	52.330	0.07
AL	1.023	0.522	2.004	1.94
AR1	0.822	0.388	1.743	1.60
AR2	1.118	0.523	2.388	1.57
CA	1.547	1.005	2.381	3.85
CO	0.530	0.192	1.465	0.93
CT1	3.182	0.643	15.758	0.39
CT2	1.465	0.376	5.699	0.54
DC	0.256	0.005	13.891	0.06
DE	7.945	0.381	165.505	0.11
FL	1.308	0.807	2.121	3.27
GA1	0.981	0.582	1.653	2.92
GA2	0.933	0.538	1.617	2.69
IA	3.408	1.204	9.651	0.89
ID	0.471	0.155	1.428	0.79
IL	0.907	0.484	1.702	2.17
IN	1.177	0.587	2.362	1.83
KY1	1.444	0.797	2.617	2.38
LA	1.385	0.642	2.988	1.54
MA1	0.386	0.164	0.911	1.27
MA2	0.534	0.155	1.831	0.65
MD1	0.583	0.195	1.745	0.81
MD2	3.778	1.213	11.761	0.76
ME1	7.408	0.359	152.841	0.11
ME2	0.987	0.282	3.453	0.63
ME3	0.769	0.210	2.818	0.59
MI	1.020	0.536	1.941	2.09
MN1	1.557	0.517	4.693	0.80
MN2	0.280	0.098	0.795	0.88
MO	1.175	0.644	2.143	2.34
MS	1.648	0.720	3.775	1.35
NC	1.556	0.792	3.054	1.93
ND	0.812	0.105	6.271	0.24
NE	2.358	0.815	6.825	0.86
NH1	0.713	0.042	11.983	0.13
NH2	1.420	0.121	16.658	0.17
NH3	0.966	0.240	3.890	0.51
NJ	1.039	0.366	2.948	0.89
NM	0.572	0.133	2.453	0.47
NV	2.541	0.499	12.942	0.38
NY1	0.394	0.208	0.748	2.10
NY2	0.877	0.488	1.576	2.43
OH1	1.598	0.940	2.718	2.84
OH2	1.282	0.769	2.138	3.01
OK	1.499	0.685	3.280	1.49
OR	1.018	0.366	2.834	0.92
PA	0.900	0.524	1.545	2.77
RI1	0.969	0.057	16.600	0.13
RI2	0.517	0.017	16.150	0.09
SC1	1.097	0.560	2.145	1.95
SC2	1.477	0.704	3.098	1.64
SD	0.814	0.169	3.910	0.41
TN	1.650	0.953	2.857	2.70
TX	1.391	0.969	1.998	4.84
UT1	1.709	0.303	9.652	0.34
UT2	2.655	0.926	7.611	0.87
UT3	2.075	0.577	7.465	0.60

```
VA1                  |  0.892      0.416     1.911          1.56
VA2                  |  0.987      0.529     1.842          2.20
VA3                  |  1.017      0.541     1.912          2.15
VA4                  |  1.026      0.506     2.080          1.78
VT                   |  0.168      0.019     1.500          0.21
WA                   |  0.940      0.428     2.064          1.48
WI                   |  0.555      0.271     1.140          1.73
WV                   |  0.466      0.166     1.311          0.90
WY                   |  4.248      0.197    91.426          0.11
Alberta              |  1.148      0.509     2.591          1.39
BC1                  |  1.478      0.683     3.196          1.53
BC2                  |  0.832      0.412     1.680          1.80
Manitoba             |  0.307      0.034     2.727          0.21
NewBrunswick         |  1.004      0.019    51.797          0.07
Newfoundland         |  0.201      0.019     2.177          0.18
NovaScotia           |  0.475      0.048     4.725          0.19
Ontario              |  1.085      0.639     1.842          2.85
PrinceEdwardI        |  1.745      0.132    23.075          0.15
Quebec               |  0.755      0.443     1.285          2.83
Saskatchewan         |  4.032      0.171    95.208          0.10
Yukon                |  2.643      0.038   181.963          0.06
---------------------+---------------------------------------------------
D+L pooled ES        |  1.083      0.978     1.199        100.00
---------------------+---------------------------------------------------

  Heterogeneity chi-squared =  86.63 (d.f. = 77) p = 0.212
  I-squared (variation in ES attributable to heterogeneity) =  11.1%
  Estimate of between-study variance Tau-squared =  0.0218

  Test of ES=1 : z=   1.53 p = 0.126
```

Figure 3.1: Forest Plot

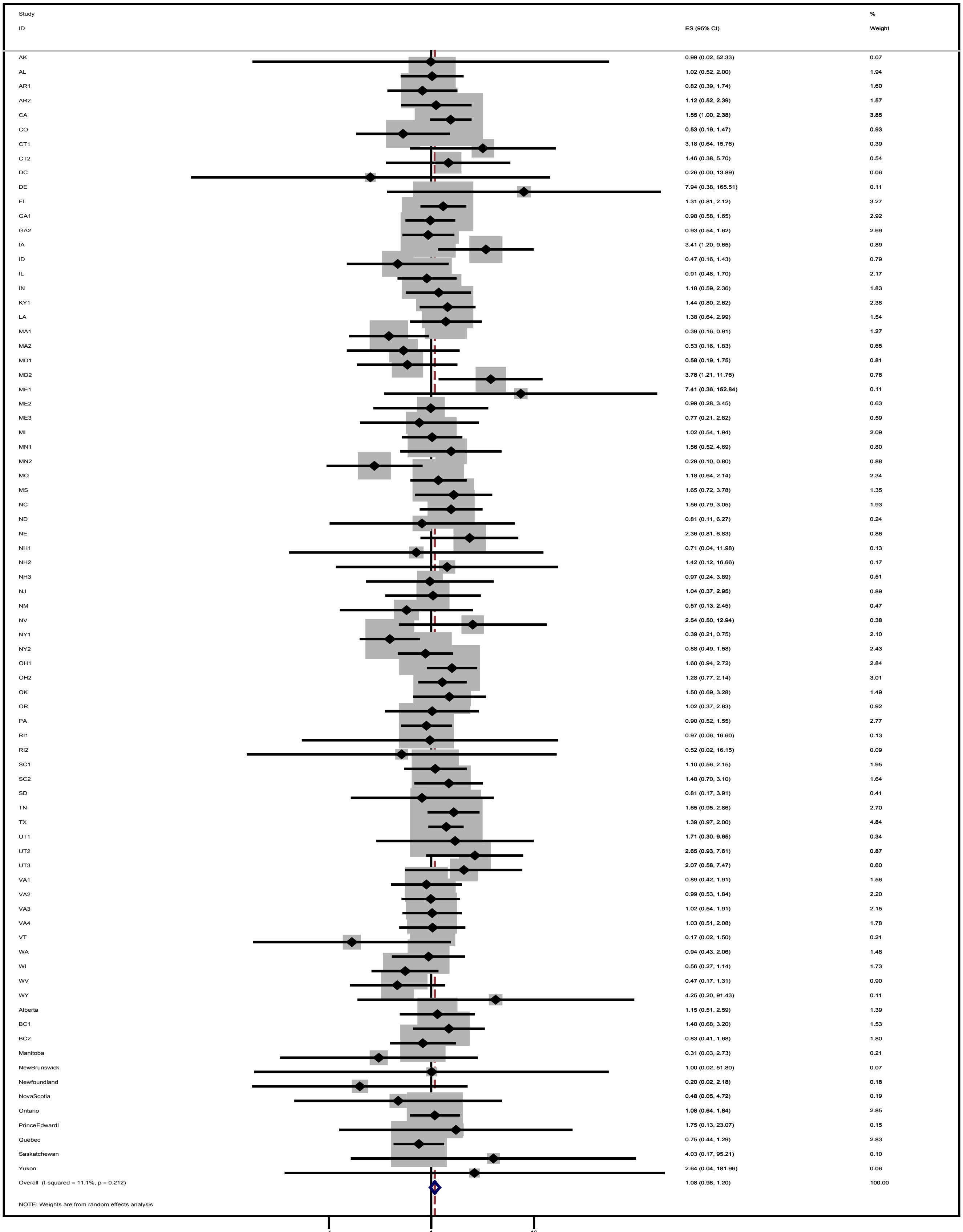
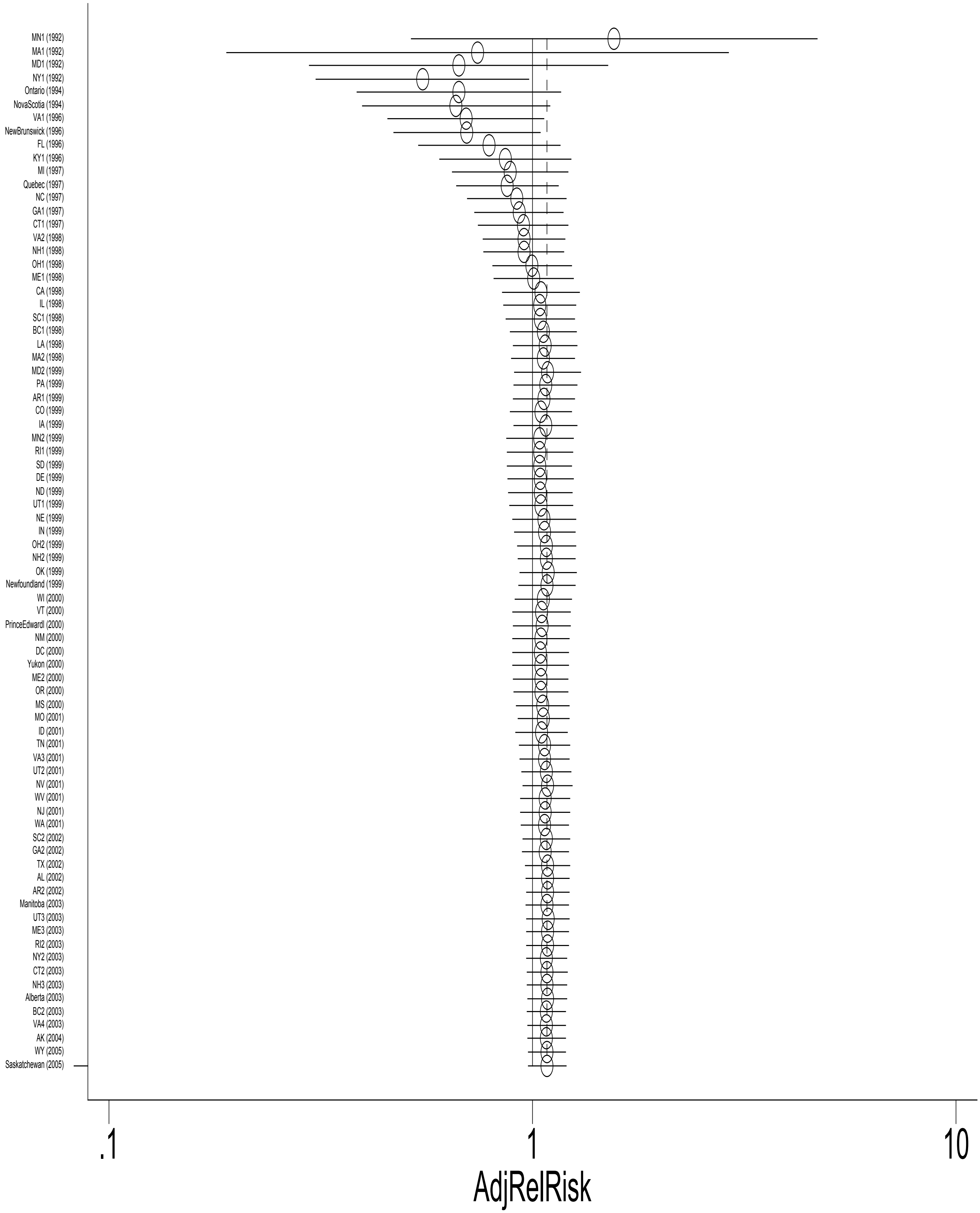


Figure 3.2: Cumulative Random Effects Meta-Analysis



Appendix 4: Summary effect for 19 year old drivers

Table 4.1: Random Effects Meta-Analysis

Study	ES	[95% Conf.	Interval]	% Weight
AK	1.834	0.058	57.727	0.08
AL	0.944	0.547	1.630	2.71
AR1	1.383	0.569	3.365	1.10
AR2	1.363	0.654	2.839	1.58
CA	1.062	0.698	1.616	4.25
CO	1.278	0.434	3.760	0.76
CT1	0.587	0.184	1.865	0.66
CT2	0.075	0.004	1.326	0.11
DC	0.233	0.004	12.664	0.06
DE	1.254	0.075	20.988	0.11
FL	1.155	0.688	1.941	2.96
GA1	1.238	0.669	2.294	2.18
GA2	1.026	0.548	1.920	2.11
IA	1.904	0.730	4.966	0.95
ID	1.075	0.330	3.499	0.64
IL	0.762	0.415	1.401	2.23
IN	0.827	0.407	1.679	1.69
KY1	0.978	0.500	1.912	1.87
LA	2.428	1.211	4.870	1.74
MA1	0.717	0.251	2.042	0.80
MA2	1.543	0.529	4.504	0.77
MD1	0.622	0.244	1.583	1.00
MD2	4.420	1.441	13.552	0.70
ME1	0.399	0.040	4.007	0.17
ME2	1.040	0.297	3.640	0.57
ME3	0.450	0.086	2.356	0.33
MI	1.365	0.709	2.628	1.95
MN1	0.322	0.122	0.850	0.93
MN2	1.029	0.454	2.336	1.28
MO	1.301	0.765	2.212	2.85
MS	0.747	0.330	1.692	1.29
NC	1.132	0.640	2.001	2.51
ND	0.389	0.033	4.602	0.15
NE	1.933	0.587	6.371	0.62
NH1	2.093	0.205	21.389	0.17
NH2	1.394	0.234	8.295	0.28
NH3	0.267	0.027	2.691	0.17
NJ	2.107	0.890	4.986	1.17
NM	1.013	0.316	3.242	0.65
NV	0.938	0.184	4.779	0.34
NY1	0.709	0.387	1.297	2.26
NY2	0.728	0.392	1.351	2.17
OH1	1.432	0.845	2.427	2.88
OH2	1.217	0.693	2.134	2.57
OK	1.939	0.839	4.482	1.23
OR	1.395	0.533	3.651	0.94
PA	0.589	0.334	1.038	2.53
RI1	0.887	0.052	15.197	0.11
RI2	1.021	0.060	17.309	0.11
SC1	0.540	0.281	1.038	1.96
SC2	1.004	0.525	1.921	1.99
SD	2.762	0.501	15.213	0.31
TN	1.869	1.007	3.468	2.17
TX	0.901	0.649	1.251	6.25
UT1	1.003	0.383	2.627	0.94
UT2	1.281	0.439	3.742	0.77
UT3	0.392	0.101	1.520	0.49

VA1	|	1.215	0.599	2.462	1.70
VA2	|	1.663	0.828	3.343	1.73
VA3	|	0.850	0.446	1.622	2.00
VA4	|	0.704	0.369	1.346	1.99
VT	|	1.018	0.060	17.180	0.11
WA	|	1.013	0.468	2.194	1.43
WI	|	1.136	0.549	2.350	1.61
WV	|	0.940	0.382	2.315	1.07
WY	|	0.176	0.007	4.202	0.09
Alberta	|	0.642	0.299	1.379	1.46
BC1	|	1.298	0.651	2.589	1.77
BC2	|	1.098	0.517	2.331	1.51
Manitoba	|	10.441	1.235	88.251	0.20
NewBrunswick	|	1.512	0.240	9.509	0.27
Newfoundland	|	0.295	0.024	3.641	0.14
NovaScotia	|	9.676	1.148	81.583	0.20
Ontario	|	0.971	0.612	1.541	3.63
PrinceEdwardI	|	0.212	0.008	5.433	0.09
Quebec	|	1.558	0.904	2.687	2.72
Saskatchewan	|	1.990	0.063	62.843	0.08
Yukon	|	2.520	0.037	173.486	0.05
D+L pooled ES	|	1.059	0.963	1.165	100.00

Heterogeneity chi-squared = 81.48 (d.f. = 77) p = 0.342
I-squared (variation in ES attributable to heterogeneity) = 5.5%
Estimate of between-study variance Tau-squared = 0.0098

Test of ES=1 : z= 1.19 p = 0.235

Figure 4.1: Forest Plot

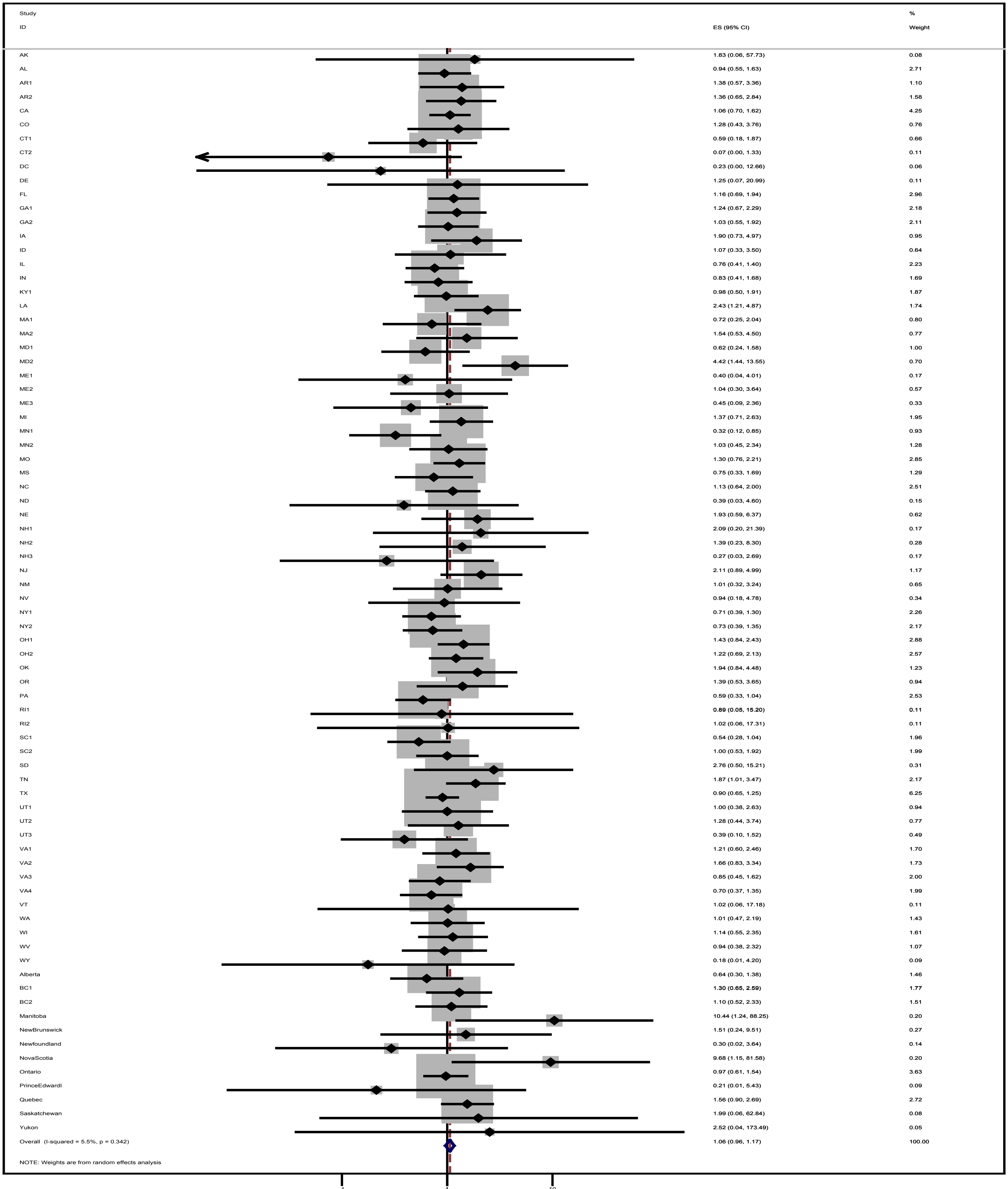
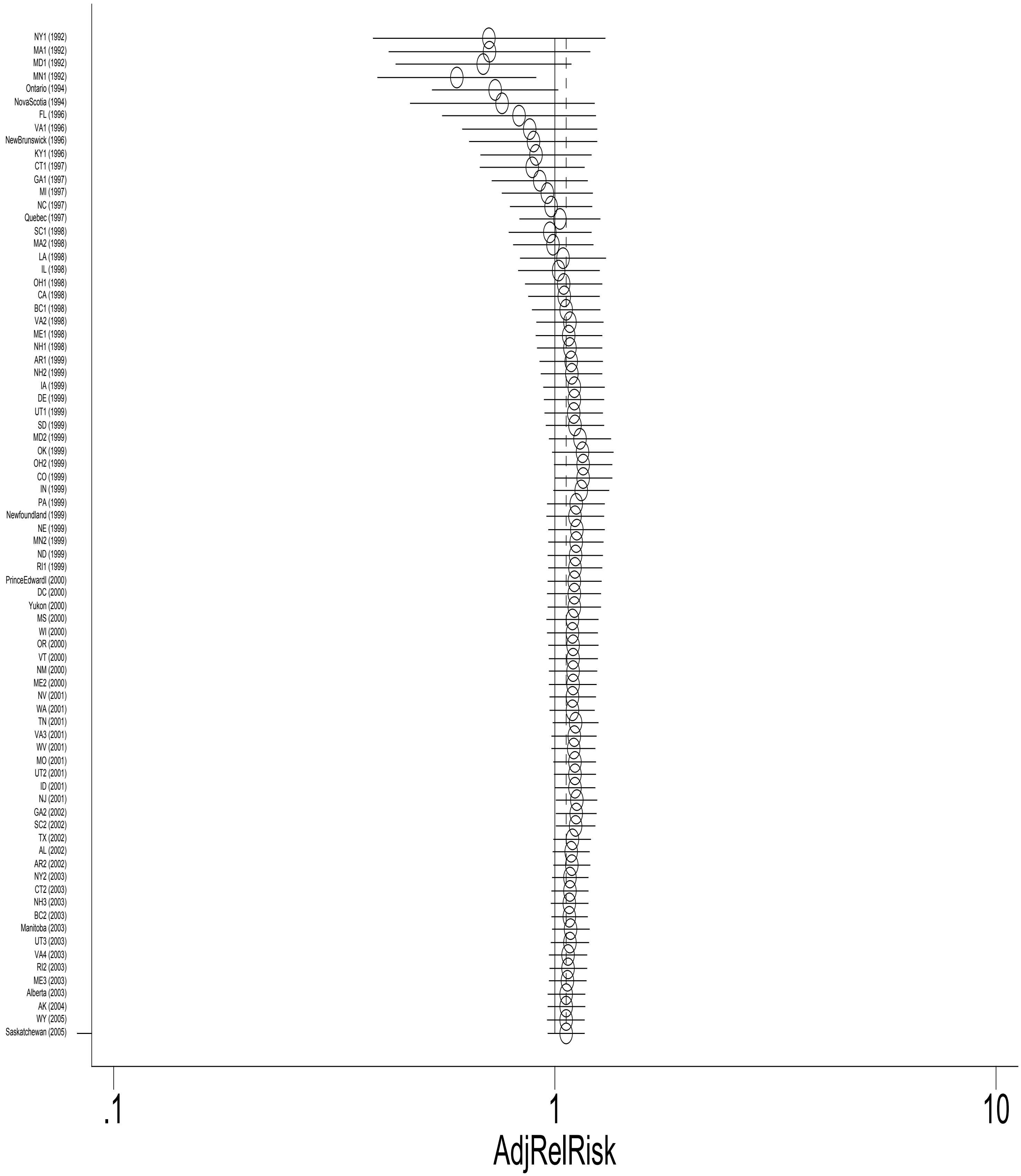


Figure 4.2: Cumulative Random Effects Meta-Analysis



Appendix 5: Meta-regression for 16 year old drivers

Table 5.1: Meta-regression

```
Meta-regression                                       Number of obs  =      48
REML estimate of between-study variance               tau2           =   .0301
% residual variation due to heterogeneity             I-squared_res  =   2.46%
Proportion of between-study variance explained        Adj R-squared  =  73.69%
Joint test for all covariates                         Model F(25,22) =    1.18
With Knapp-Hartung modification                      Prob > F       =  0.3466
```

logadjRR7	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
year	-.1074513	.0644252	-1.67	0.110	-.2410609	.0261584
_Iiihs_2	.2196012	.3505599	0.63	0.537	-.5074155	.946618
_Iiihs_3	-.3659963	.4823285	-0.76	0.456	-1.366284	.6342918
lslengthmin	.170582	.0852881	2.00	0.058	-.0062947	.3474587
lslengthmax	-.0454361	.065892	-0.69	0.498	-.1820876	.0912155
lspar1min	-.0172798	.0097601	-1.77	0.091	-.0375211	.0029614
_Ilspar2re~1	.2156036	.4304987	0.50	0.621	-.6771961	1.108403
lsnightrl	-.0020125	.0600872	-0.03	0.974	-.1266257	.1226007
lsnightifsup	.5086947	.6584168	0.77	0.448	-.8567781	1.874168
lspas1rec	.6685581	1.296098	0.52	0.611	-2.019386	3.356502
lsiffam	-.1769388	1.641356	-0.11	0.915	-3.580903	3.227025
lsifdrivered	2.59863	2.42896	1.07	0.296	-2.438725	7.635985
lsage	.5202102	.416306	1.25	0.225	-.3431557	1.383576
lsredu	-.6997212	.4169904	-1.68	0.107	-1.564506	.165064
lsexit	1.778373	1.589924	1.12	0.275	-1.518927	5.075673
_Ilsdredre~1	-.2630512	.2516453	-1.05	0.307	-.7849315	.2588292
_Ilsdredre~2	.4599831	1.870905	0.25	0.808	-3.420036	4.340002
isnightrllib	-.0955647	.0806419	-1.19	0.249	-.2628058	.0716763
isnightifwor	1.290175	2.80328	0.46	0.650	-4.523471	7.103821
ispas1rec	-2.160185	.8043738	-2.69	0.014	-3.828354	-.4920157
isifsup	.6340943	.9821616	0.65	0.525	-1.402784	2.670973
isiffam	2.114694	.7944557	2.66	0.014	.4670933	3.762294
isage	-1.211882	.8049229	-1.51	0.146	-2.88119	.4574256
_Iisdred_2	-.322202	.5473806	-0.59	0.562	-1.4574	.8129958
isexit	-3.9084	2.041057	-1.91	0.069	-8.141293	.3244924
_cons	226.3739	129.2742	1.75	0.094	-41.72433	494.4721

Figure 5.1: Full Bayesian Meta-regression

BAYES, GIBBS SAMPLING, RANDOM MODEL, 16 YEAR OLDS (50,000 iterations, burn-in period of 1000 iterations), RANDOM SEED 1

$$\mathrm{logadjRR7}_{ij} \sim \mathrm{N}(XB, \Omega)$$

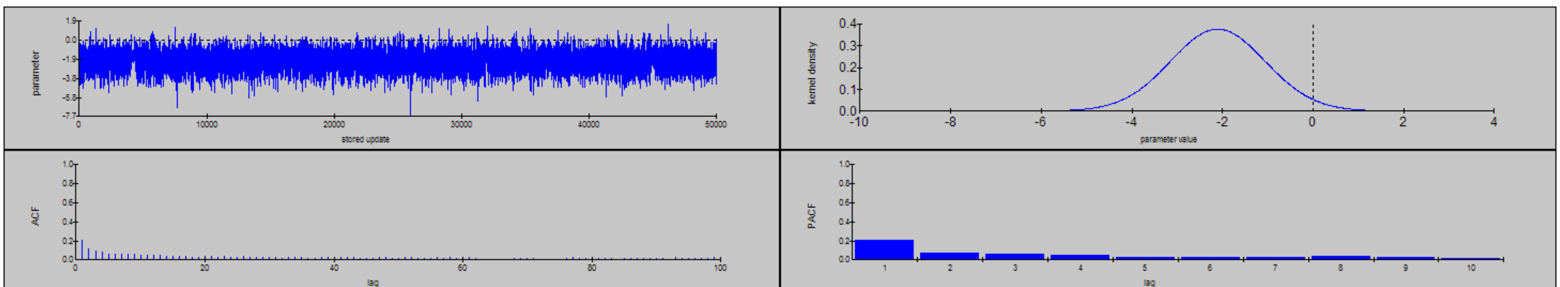
$$\begin{aligned}\mathrm{logadjRR7}_{ij} = {} & \beta_{0j}\mathrm{CONS} + \text{-}0.100(0.073)\mathrm{year}_j + 0.225(0.377)\mathrm{g}_j + \text{-}0.408(0.532)\mathrm{m}_j + 0.164(0.093)\mathrm{lslengthmin}_j + \text{-}0.048(0.069)\mathrm{lslengthmax}_j + \\ & \text{-}0.017(0.010)\mathrm{lspar1min}_j + 0.218(0.456)\mathrm{lspar2rec\_1}_j + \text{-}0.007(0.068)\mathrm{lsnightrl}_j + 0.577(0.717)\mathrm{lsnightifsup\_1}_j + \\ & 0.638(1.344)\mathrm{lspas1rec\_1}_j + \text{-}0.132(1.712)\mathrm{lsiffam\_1}_j + 2.576(2.518)\mathrm{lsifdrivered\_1}_j + 0.502(0.455)\mathrm{lsage}_j + \text{-}0.693(0.432)\mathrm{lsredu}_j + \\ & 1.836(1.659)\mathrm{lsexit\_1}_j + \text{-}0.269(0.276)\mathrm{lsdredrec\_1}_j + 0.476(1.935)\mathrm{lsdredrec\_2}_j + \text{-}0.096(0.088)\mathrm{isnightrllib}_j + \\ & 1.379(2.941)\mathrm{isnightifwor\_1}_j + \text{-}2.102(0.877)\mathrm{ispas1rec\_1}_j + 0.574(1.058)\mathrm{isifsup\_1}_j + 2.011(0.888)\mathrm{isiffam\_1}_j + \text{-}1.210(0.861)\mathrm{isage}_j + \\ & \text{-}0.363(0.608)\mathrm{isdred\_2}_j + \text{-}3.962(2.158)\mathrm{isexit\_1}_j + e_{1ij}\mathrm{logadjRR7se}_j\end{aligned}$$

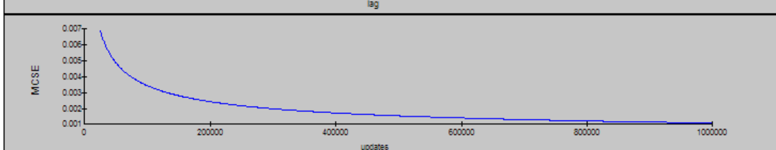
$$\beta_{0j} = 211.074(146.598) + u_{0j}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 0.078(0.118) \end{bmatrix}$$

$$\begin{bmatrix} e_{1ij} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_e)\ :\ \Omega_e = \begin{bmatrix} 1.037(0.473) \end{bmatrix}$$

Deviance(MCMC) = 98.096(48 of 78 cases in use)





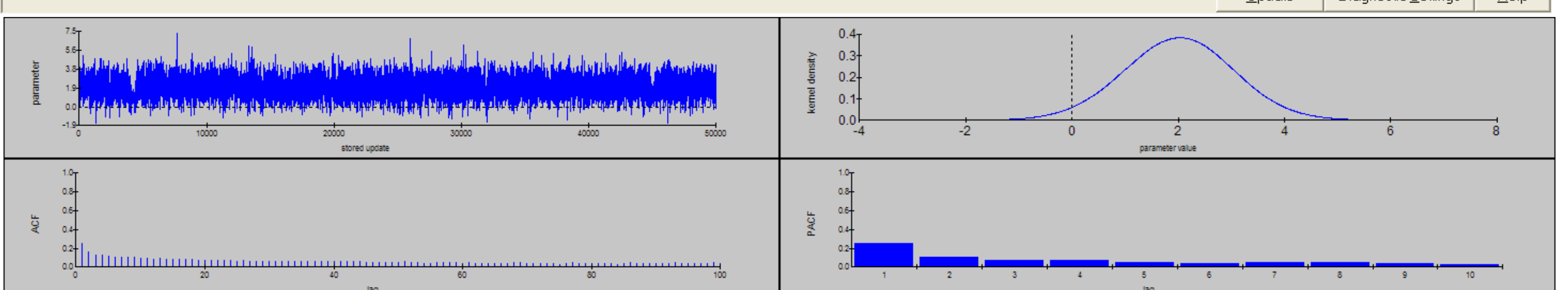
Accuracy Diagnostics

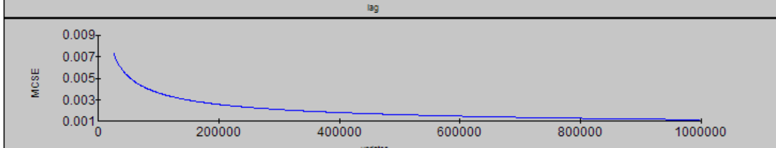
Raftery-Lewis (quantile) : Nhat = (4090, 5301)
when q = (.025, .975), r = .005 and s = .95
Brooks-Draper (mean) : Nhat = 1817
when k = 2 sigfigs and alpha = .05

Summary Statistics

param name : β_{21} posterior mean = -2.102 (0.0049) SD = 0.877 mode = -2.107
quantiles : 2.5% = -3.833, 5% = -3.540, 50% = -2.104, 95% = -0.659, 97.5% = -0.364
50000 actual iterations storing every iteration. Effective Sample Size (ESS) = 23295.

Update | Diagnostic Settings | Help





Accuracy Diagnostics

Raftery-Lewis (quantile) : Nhat = (5733, 4090)
when q = (.025, .975), r = .005 and s = .95
Brooks-Draper (mean) : Nhat = 2040
when k = 2 sigfigs and alpha = .05

Summary Statistics

param name : β_{23} posterior mean = 2.011 (0.0052) SD = 0.888 mode = 2.032
quantiles : 2.5% = 0.237, 5% = 0.541, 50% = 2.020, 95% = 3.457, 97.5% = 3.762
50000 actual iterations storing every iteration. Effective Sample Size (ESS) = 14616.

Update | Diagnostic Settings | Help

BAYES, GIBSS SAMPLING, RANDOM MODEL, 16 YEAR OLDS (50,000 iterations, burn-in period of 1,000 iterations), RANDOM SEED 2

$\text{logadjRR7}_{ij} \sim \text{N}(XB, \Omega)$

$\text{logadjRR7}_{ij} = \beta_{0j}\text{CONS} + -0.098(0.073)\text{year}_j + 0.225(0.376)\text{g}_j + -0.412(0.525)\text{m}_j + 0.163(0.092)\text{lslengthmin}_j + -0.048(0.069)\text{lslengthmax}_j + -0.017(0.010)\text{lspar1min}_j + 0.221(0.454)\text{lspar2rec\_1}_j + -0.007(0.068)\text{lsnightrl}_j + 0.573(0.717)\text{lsnightifsup\_1}_j + 0.640(1.335)\text{lspas1rec\_1}_j + -0.134(1.709)\text{lsiffam\_1}_j + 2.577(2.477)\text{lsifdrivered\_1}_j + 0.498(0.450)\text{lsage}_j + -0.690(0.429)\text{lsredu}_j + 1.831(1.647)\text{lsexit\_1}_j + -0.266(0.273)\text{lsdredrec\_1}_j + 0.472(1.926)\text{lsdredrec\_2}_j + -0.096(0.088)\text{isnightrllib}_j + 1.400(2.927)\text{isnightifwor\_1}_j + -2.095(0.869)\text{ispas1rec\_1}_j + 0.566(1.050)\text{isifsup\_1}_j + 2.000(0.880)\text{isiffam\_1}_j + -1.207(0.858)\text{isage}_j + -0.369(0.606)\text{isdred\_2}_j + -3.976(2.157)\text{isexit\_1}_j + e_{1ij}\text{logadjRR7se}_j$

$\beta_{0j} = 207.961(145.489) + u_{0j}$

$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.080(0.120)\right]$

$\left[e_{1ij}\right] \sim \text{N}(0,\ \Omega_e)\ :\ \Omega_e = \left[1.019(0.477)\right]$

Deviance(MCMC) = 97.318(48 of 78 cases in use)

BAYES, GIBSS SAMPLING, RANDOM MODEL, 16 YEAR OLDS (50,000 iterations, burn-in period of 1,000 iterations), RANDOM SEED 3

$\text{logadjRR7}_{ij} \sim \text{N}(XB, \Omega)$

$\text{logadjRR7}_{ij} = \beta_{0j}\text{CONS} + -0.099(0.074)\text{year}_j + 0.223(0.381)\text{g}_j + -0.403(0.537)\text{m}_j + 0.163(0.093)\text{lslengthmin}_j + -0.047(0.069)\text{lslengthmax}_j + -0.017(0.011)\text{lspar1min}_j + 0.222(0.456)\text{lspar2rec\_1}_j + -0.007(0.068)\text{lsnightrl}_j + 0.570(0.720)\text{lsnightifsup\_1}_j + 0.631(1.345)\text{lspas1rec\_1}_j + -0.130(1.710)\text{lsiffam\_1}_j + 2.605(2.534)\text{lsifdrivered\_1}_j + 0.500(0.454)\text{lsage}_j + -0.692(0.434)\text{lsredu}_j + 1.816(1.663)\text{lsexit\_1}_j + -0.270(0.274)\text{lsdredrec\_1}_j + 0.487(1.942)\text{lsdredrec\_2}_j + -0.096(0.088)\text{isnightrllib}_j + 1.362(2.961)\text{isnightifwor\_1}_j + -2.095(0.882)\text{ispas1rec\_1}_j + 0.569(1.066)\text{isifsup\_1}_j + 2.003(0.898)\text{isiffam\_1}_j + -1.201(0.863)\text{isage}_j + -0.363(0.610)\text{isdred\_2}_j + -3.955(2.182)\text{isexit\_1}_j + e_{1ij}\text{logadjRR7se}_j$

$\beta_{0j} = 210.169(148.128) + u_{0j}$

$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.081(0.128)\right]$

$\left[e_{1ij}\right] \sim \text{N}(0,\ \Omega_e)\ :\ \Omega_e = \left[1.038(0.481)\right]$

Deviance(MCMC) = 97.334(48 of 78 cases in use)

BAYES, GIBSS SAMPLING, RANDOM MODEL, 16 YEAR OLDS (50,000 iterations, burn-in period of 1,000 iterations), RANDOM SEED 4

$$\mathrm{logadjRR7}_{ij} \sim \mathrm{N}(XB, \Omega)$$

$$\begin{aligned}\mathrm{logadjRR7}_{ij} = {} & \beta_{0j}\mathrm{CONS} + -0.095(0.076)\mathrm{year}_j + 0.225(0.381)\mathrm{g}_j + -0.414(0.538)\mathrm{m}_j + 0.163(0.094)\mathrm{lslengthmin}_j + -0.048(0.068)\mathrm{lslengthmax}_j + \\ & -0.017(0.010)\mathrm{lspar1min}_j + 0.217(0.458)\mathrm{lspar2rec\_1}_j + -0.009(0.069)\mathrm{lsnightrl}_j + 0.592(0.718)\mathrm{lsnightifsup\_1}_j + \\ & 0.625(1.330)\mathrm{lspas1rec\_1}_j + -0.115(1.699)\mathrm{lsiffam\_1}_j + 2.558(2.501)\mathrm{lsifdrivered\_1}_j + 0.493(0.453)\mathrm{lsage}_j + -0.689(0.427)\mathrm{lsredu}_j + \\ & 1.838(1.647)\mathrm{lsexit\_1}_j + -0.273(0.279)\mathrm{lsdredrec\_1}_j + 0.476(1.918)\mathrm{lsdredrec\_2}_j + -0.096(0.089)\mathrm{isnightrllib}_j + \\ & 1.442(2.929)\mathrm{isnightifwor\_1}_j + -2.071(0.884)\mathrm{ispas1rec\_1}_j + 0.554(1.062)\mathrm{isifsup\_1}_j + 1.968(0.910)\mathrm{isiffam\_1}_j + -1.199(0.869)\mathrm{isage}_j + \\ & -0.366(0.608)\mathrm{isdred\_2}_j + -3.994(2.167)\mathrm{isexit\_1}_j + e_{1ij}\mathrm{logadjRR7se}_j\end{aligned}$$

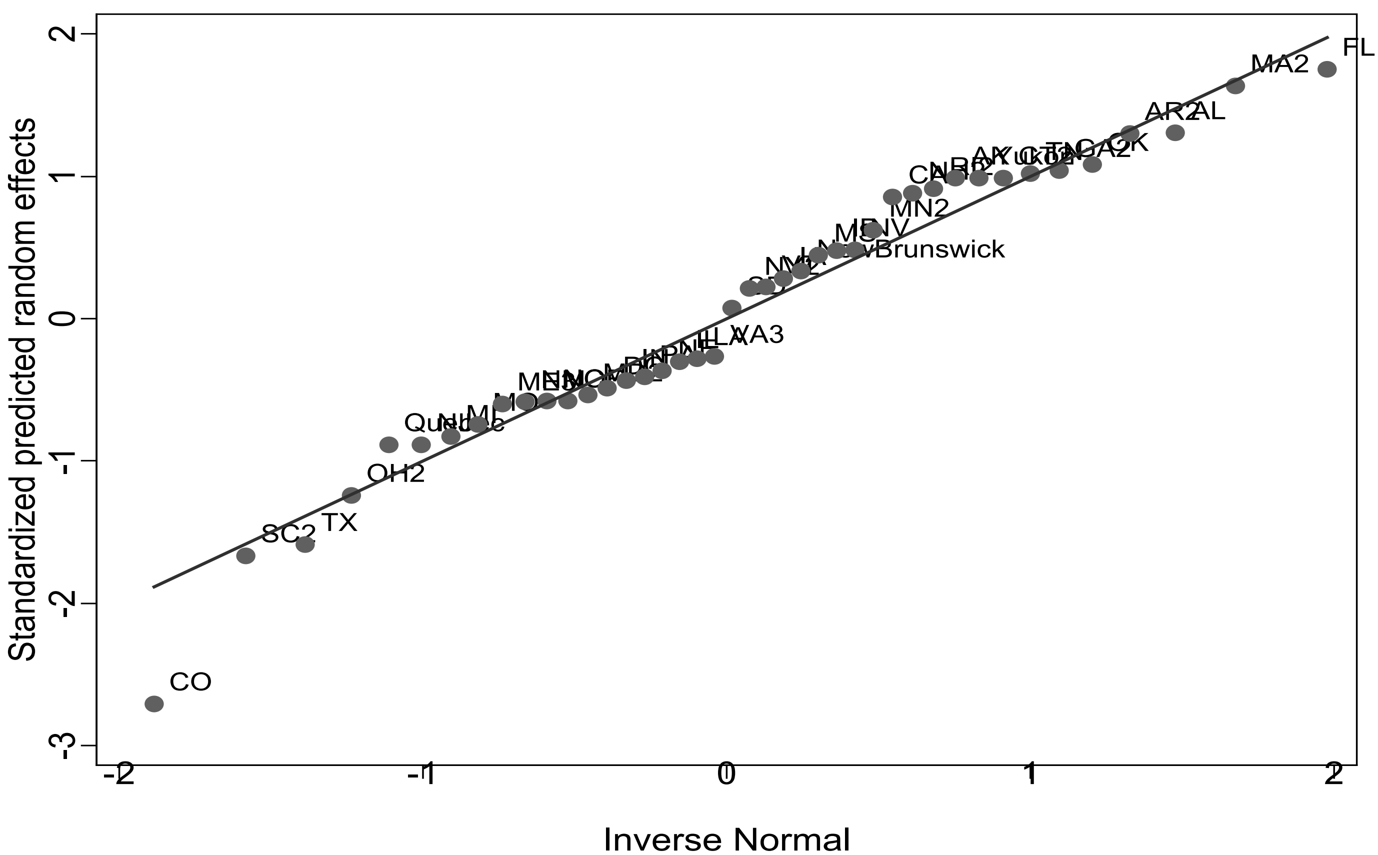
$$\beta_{0j} = 202.522(151.786) + u_{0j}$$

$$\left[u_{0j}\right] \sim \mathrm{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.093(0.140)\right]$$

$$\left[e_{1ij}\right] \sim \mathrm{N}(0,\ \Omega_e)\ :\ \Omega_e = \left[1.007(0.505)\right]$$

Deviance(MCMC) = 93.962(48 of 78 cases in use)

Figure 5.2: Normal probability plot of standardized shrunken residuals



Appendix 6: Meta-regression for 18 year old drivers

Table 6.1: Meta-regression

```
Meta-regression                                        Number of obs  =      48
REML estimate of between-study variance                tau2           =       0
% residual variation due to heterogeneity              I-squared_res  =  11.43%
Proportion of between-study variance explained         Adj R-squared  =      .%
Joint test for all covariates                          Model F(25,22) =    0.87
With Knapp-Hartung modification                       Prob > F       =  0.6351
------------------------------------------------------------------------------
   logadjRR9 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        year |   .0108451   .0527708     0.21   0.839    -.0985949    .1202851
     _Iiihs_2 |   .0730795   .2698139     0.27   0.789    -.4864803    .6326394
     _Iiihs_3 |  -.2167965   .3504076    -0.62   0.542    -.9434974    .5099044
  lslengthmin |   .0286035   .0680428     0.42   0.678    -.1125086    .1697156
  lslengthmax |   -.055624   .0538544    -1.03   0.313    -.1673112    .0560632
    lspar1min |  -.0082214   .0070555    -1.17   0.256    -.0228536    .0064109
 _Ilspar2re~1 |   .4045334   .2982076     1.36   0.189    -.2139113    1.022978
    lsnightrl |   .0426433   .0478848     0.89   0.383    -.0566638    .1419503
 lsnightifsup |  -.1207104   .4950782    -0.24   0.810     -1.14744     .906019
    lspas1rec |  -.1823846   .6793941    -0.27   0.791    -1.591362    1.226593
      lsiffam |   2.119684   1.775762     1.19   0.245    -1.563021    5.802389
 lsifdrivered |   .2768912   2.012674     0.14   0.892    -3.897139    4.450921
        lsage |   .1051972   .3343059     0.31   0.756    -.5881109    .7985052
       lsredu |  -.2408143   .3354282    -0.72   0.480    -.9364498    .4548212
       lsexit |   .2461586   1.209511     0.20   0.841    -2.262214    2.754531
 _Ilsdredre~1 |  -.4232671   .1889674    -2.24   0.036    -.8151615   -.0313728
 _Ilsdredre~2 |   .4547635   1.101381     0.41   0.684    -1.829361    2.738888
 isnightrllib |   -.018547   .0615327    -0.30   0.766    -.1461581    .1090641
 isnightifwor |  -.4938242   1.961843    -0.25   0.804    -4.562438    3.574789
    ispas1rec |   .0643973   .6731866     0.10   0.925    -1.331706    1.460501
      isifsup |   .6135038   .8539684     0.72   0.480    -1.157518    2.384526
      isiffam |  -.3337816   .6759174    -0.49   0.626    -1.735548    1.067985
        isage |  -.3207081   .5763126    -0.56   0.583    -1.515907    .8744911
   _Iisdred_2 |   .1629145   .4435683     0.37   0.717    -.7569899    1.082819
       isexit |  -.6391904   1.569569    -0.41   0.688    -3.894278    2.615897
        _cons |  -17.42221   105.2834    -0.17   0.870    -235.7666    200.9222
------------------------------------------------------------------------------
```

Figure 6.1: Full Bayesian Meta-regression

BAYES, GIBBS SAMPLING, RANDOM MODEL, 18 YEAR OLDS (50,000 iterations, burn-in period of 1000 iterations), RANDOM SEED 1

$$\mathrm{logadjRR9}_{ij} \sim \mathrm{N}(XB, \Omega)$$

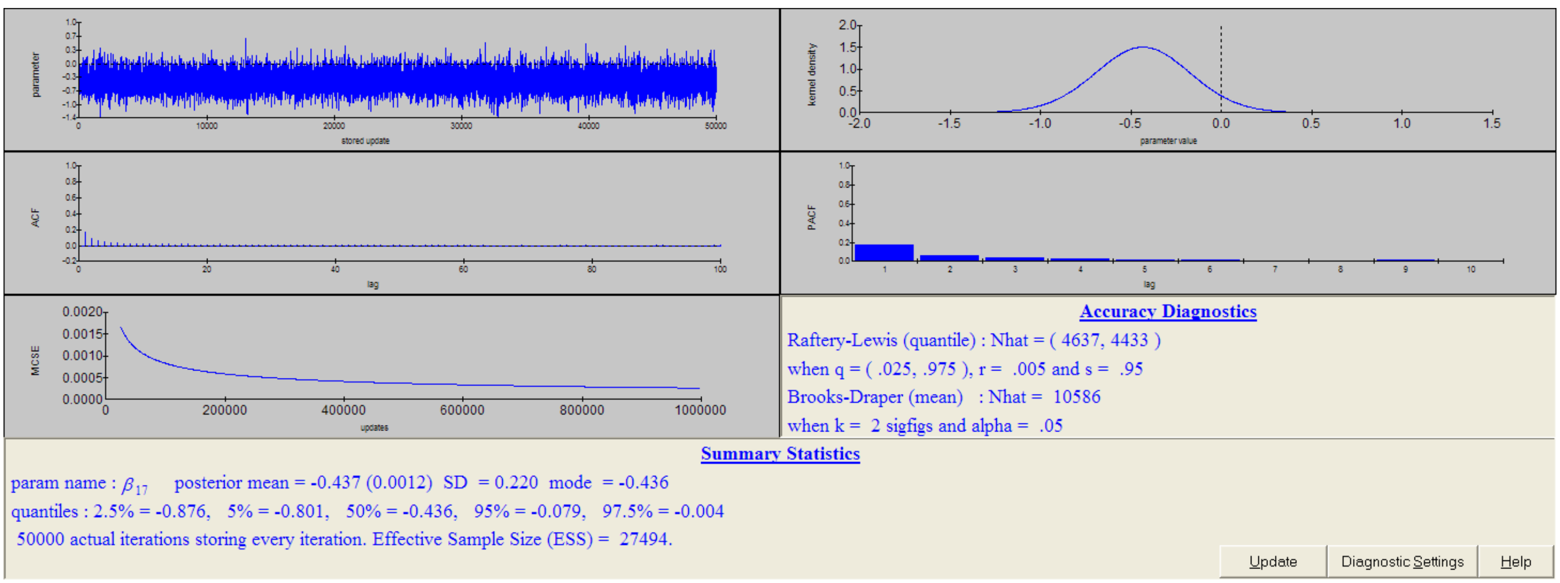
$$\begin{aligned}\mathrm{logadjRR9}_{ij} = {} & \beta_{0j}\mathrm{CONS} + 0.007(0.063)\mathrm{year}_j + 0.100(0.318)\mathrm{g}_j + -0.190(0.416)\mathrm{m}_j + 0.023(0.079)\mathrm{lslengthmin}_j + -0.054(0.059)\mathrm{lslengthmax}_j + \\ & -0.008(0.008)\mathrm{lspar1min}_j + 0.385(0.339)\mathrm{lspar2rec\_1}_j + 0.037(0.058)\mathrm{lsnightrl}_j + -0.108(0.561)\mathrm{lsnightifsup\_1}_j + \\ & -0.155(0.762)\mathrm{lspas1rec\_1}_j + 2.056(1.907)\mathrm{lsiffam\_1}_j + 0.226(2.177)\mathrm{lsifdrivered\_1}_j + 0.049(0.394)\mathrm{lsage}_j + -0.231(0.368)\mathrm{lsredu}_j + \\ & 0.199(1.342)\mathrm{lsexit\_1}_j + -0.437(0.220)\mathrm{lsdredrec\_1}_j + 0.528(1.213)\mathrm{lsdredrec\_2}_j + -0.008(0.074)\mathrm{isnightrllib}_j + \\ & -0.616(2.211)\mathrm{isnightifwor\_1}_j + 0.058(0.762)\mathrm{ispas1rec\_1}_j + 0.683(0.958)\mathrm{isifsup\_1}_j + -0.340(0.761)\mathrm{isiffam\_1}_j + -0.249(0.673)\mathrm{isage}_j + \\ & 0.156(0.512)\mathrm{isdred\_2}_j + -0.520(1.818)\mathrm{isexit\_1}_j + e_{1ij}\mathrm{logadjRR9se}_j\end{aligned}$$

$$\beta_{0j} = -9.833(124.941) + u_{0j}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 0.040(0.074) \end{bmatrix}$$

$$\begin{bmatrix} e_{1ij} \end{bmatrix} \sim \mathrm{N}(0,\ \Omega_e)\ :\ \Omega_e = \begin{bmatrix} 1.265(0.534) \end{bmatrix}$$

Deviance(MCMC) = 84.436(48 of 78 cases in use)



BAYES, GIBSS SAMPLING, RANDOM MODEL, 18 YEAR OLDS (50,000 iterations, burn-in period of 1,000 iterations), RANDOM SEED 2

$\text{logadjRR9}_{ij} \sim \text{N}(XB, \Omega)$

$\text{logadjRR9}_{ij} = \beta_{0j}\text{CONS} + 0.009(0.061)\text{year}_j + 0.095(0.311)\text{g}_j + -0.199(0.406)\text{m}_j + 0.023(0.078)\text{lslengthmin}_j + -0.054(0.059)\text{lslengthmax}_j + -0.008(0.008)\text{lspar1min}_j + 0.387(0.339)\text{lspar2rec\_1}_j + 0.038(0.057)\text{lsnightrl}_j + -0.114(0.560)\text{lsnightifsup\_1}_j + -0.163(0.759)\text{lspas1rec\_1}_j + 2.072(1.913)\text{lsiffam\_1}_j + 0.190(2.170)\text{lsifdrivered\_1}_j + 0.055(0.389)\text{lsage}_j + -0.226(0.368)\text{lsredu}_j + 0.198(1.345)\text{lsexit\_1}_j + -0.435(0.219)\text{lsdredrec\_1}_j + 0.509(1.221)\text{lsdredrec\_2}_j + -0.009(0.074)\text{isnightrllib}_j + -0.584(2.177)\text{isnightifwor\_1}_j + 0.060(0.760)\text{ispas1rec\_1}_j + 0.665(0.952)\text{isifsup\_1}_j + -0.340(0.760)\text{isiffam\_1}_j + -0.250(0.670)\text{isage}_j + 0.163(0.504)\text{isdred\_2}_j + -0.529(1.795)\text{isexit\_1}_j + e_{1ij}\text{logadjRR9se}_j$

$\beta_{0j} = -13.980(121.271) + u_{0j}$

$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.035(0.064)\right]$

$\left[e_{1ij}\right] \sim \text{N}(0,\ \Omega_e)\ :\ \Omega_e = \left[1.270(0.503)\right]$

Deviance(MCMC) = 85.632(48 of 78 cases in use)

BAYES, GIBSS SAMPLING, RANDOM MODEL, 18 YEAR OLDS (50,000 iterations, burn-in period of 1,000 iterations), RANDOM SEED 3

$\text{logadjRR9}_{ij} \sim \text{N}(XB, \Omega)$

$\text{logadjRR9}_{ij} = \beta_{0j}\text{CONS} + 0.007(0.062)\text{year}_j + 0.102(0.316)\text{g}_j + -0.190(0.413)\text{m}_j + 0.022(0.078)\text{lslengthmin}_j + -0.054(0.059)\text{lslengthmax}_j + -0.007(0.008)\text{lspar1min}_j + 0.379(0.342)\text{lspar2rec\_1}_j + 0.037(0.058)\text{lsnightrl}_j + -0.105(0.564)\text{lsnightifsup\_1}_j + -0.152(0.761)\text{lspas1rec\_1}_j + 2.063(1.892)\text{lsiffam\_1}_j + 0.185(2.173)\text{lsifdrivered\_1}_j + 0.035(0.403)\text{lsage}_j + -0.227(0.368)\text{lsredu}_j + 0.193(1.338)\text{lsexit\_1}_j + -0.441(0.222)\text{lsdredrec\_1}_j + 0.551(1.228)\text{lsdredrec\_2}_j + -0.005(0.076)\text{isnightrllib}_j + -0.641(2.189)\text{isnightifwor\_1}_j + 0.059(0.762)\text{ispas1rec\_1}_j + 0.685(0.954)\text{isifsup\_1}_j + -0.340(0.760)\text{isiffam\_1}_j + -0.229(0.681)\text{isage}_j + 0.163(0.512)\text{isdred\_2}_j + -0.498(1.795)\text{isexit\_1}_j + e_{1ij}\text{logadjRR9se}_j$

$\beta_{0j} = -10.417(123.904) + u_{0j}$

$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.047(0.083)\right]$

$\left[e_{1ij}\right] \sim \text{N}(0,\ \Omega_e)\ :\ \Omega_e = \left[1.238(0.535)\right]$

Deviance(MCMC) = 83.332(48 of 78 cases in use)

BAYES, GIBSS SAMPLING, RANDOM MODEL, 18 YEAR OLDS (50,000 iterations, burn-in period of 1,000 iterations), RANDOM SEED 4

$$\text{logadjRR9}_{ij} \sim \text{N}(XB, \Omega)$$

$$\begin{aligned}\text{logadjRR9}_{ij} = {} & \beta_{0j}\text{CONS} + 0.008(0.062)\text{year}_j + 0.102(0.314)\text{g}_j + -0.196(0.412)\text{m}_j + 0.024(0.079)\text{lslengthmin}_j + -0.054(0.059)\text{lslengthmax}_j + \\ & -0.008(0.008)\text{lspar1min}_j + 0.382(0.342)\text{lspar2rec\_1}_j + 0.038(0.058)\text{lsnightrl}_j + -0.108(0.567)\text{lsnightifsup\_1}_j + \\ & -0.165(0.761)\text{lspas1rec\_1}_j + 2.059(1.906)\text{lsiffam\_1}_j + 0.206(2.202)\text{lsifdrivered\_1}_j + 0.044(0.399)\text{lsage}_j + -0.228(0.370)\text{lsredu}_j + \\ & 0.206(1.355)\text{lsexit\_1}_j + -0.437(0.222)\text{lsdredrec\_1}_j + 0.519(1.231)\text{lsdredrec\_2}_j + -0.007(0.075)\text{isnightrllib}_j + \\ & -0.574(2.184)\text{isnightifwor\_1}_j + 0.051(0.765)\text{ispas1rec\_1}_j + 0.684(0.964)\text{isifsup\_1}_j + -0.336(0.765)\text{isiffam\_1}_j + -0.238(0.681)\text{isage}_j + \\ & 0.156(0.512)\text{isdred\_2}_j + -0.535(1.798)\text{isexit\_1}_j + e_{1ij}\text{logadjRR9se}_j\end{aligned}$$

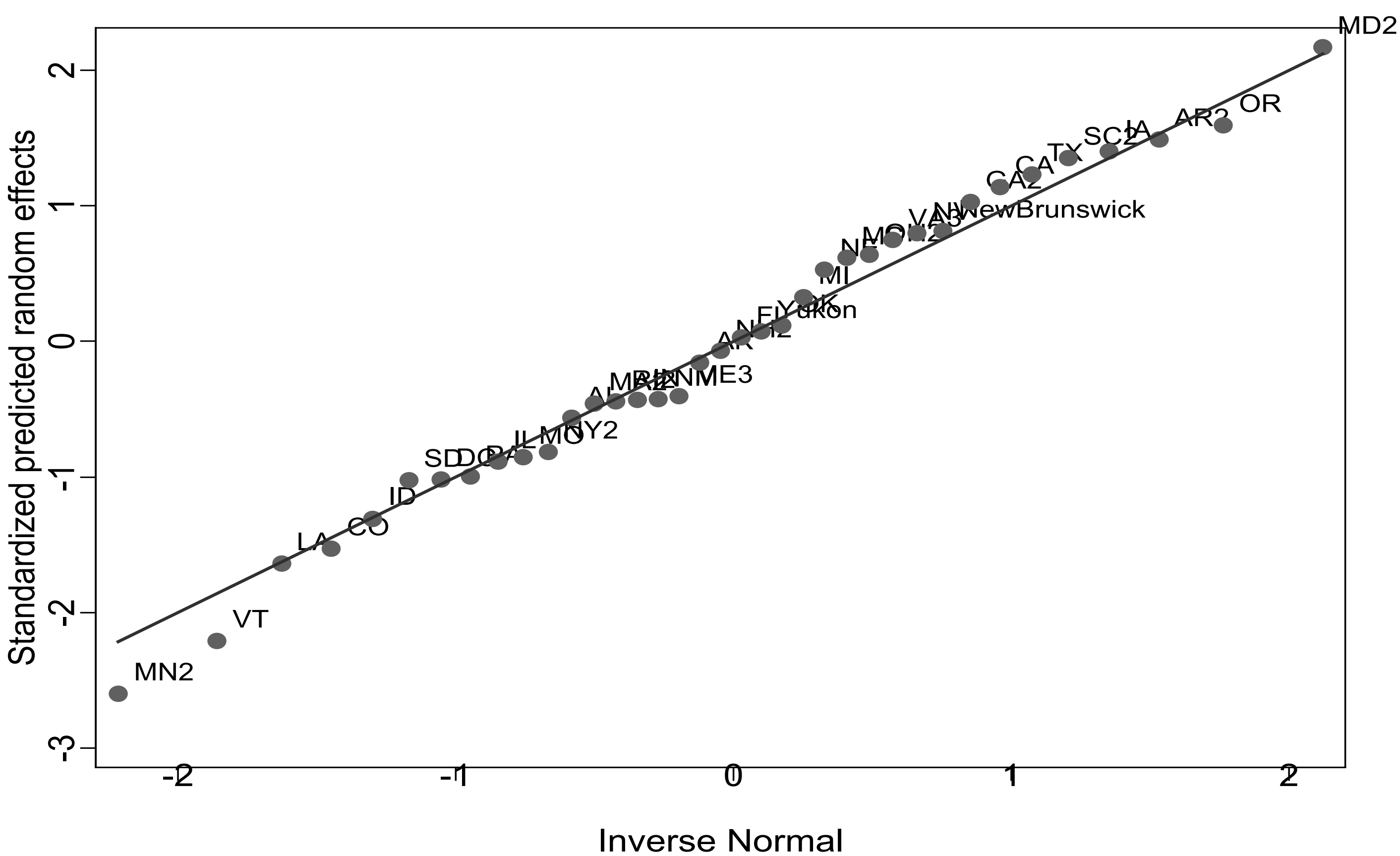
$$\beta_{0j} = -12.145(123.415) + u_{0j}$$

$$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.041(0.072)\right]$$

$$\left[e_{1ij}\right] \sim \text{N}(0,\ \Omega_e)\ :\ \Omega_e = \left[1.271(0.548)\right]$$

Deviance(MCMC) = 84.931(48 of 78 cases in use)

Figure 6.2: Normal probability plot of standardized shrunken residuals



Appendix 7: Meta-regression for 19 year old drivers

Table 7.1: Meta-regression

Meta-regression	Number of obs	=	47
REML estimate of between-study variance	tau2	=	0
% residual variation due to heterogeneity	I-squared_res	=	0.00%
Proportion of between-study variance explained	Adj R-squared	=	100.00%
Joint test for all covariates	Model F(25,21)	=	1.53
With Knapp-Hartung modification	Prob > F	=	0.1630

logadjRR10	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
year	-.0477804	.0491847	-0.97	0.342	-.1500655	.0545047
_Iiihs_2	.1787797	.2516661	0.71	0.485	-.3445887	.7021481
_Iiihs_3	-.12238	.3198451	-0.38	0.706	-.7875344	.5427744
lslengthmin	-.0325761	.0668439	-0.49	0.631	-.1715856	.1064335
lslengthmax	.0362948	.0530123	0.68	0.501	-.0739503	.1465399
lspar1min	-.005056	.0067624	-0.75	0.463	-.0191193	.0090073
_Ilspar2re~1	.0971577	.2767202	0.35	0.729	-.4783135	.672629
lsnightrl	.103704	.0468253	2.21	0.038	.0063254	.2010826
lsnightifsup	-.1092041	.443524	-0.25	0.808	-1.031563	.8131545
lspas1rec	.2855166	.599157	0.48	0.639	-.9604986	1.531532
lsiffam	.3092276	1.539185	0.20	0.843	-2.891682	3.510137
lsifdrivered	-2.581543	1.863311	-1.39	0.180	-6.456511	1.293425
lsage	-.0221124	.3083675	-0.07	0.944	-.6633977	.6191728
lsredu	-.0666995	.260502	-0.26	0.800	-.6084432	.4750441
lsexit	2.586603	1.37388	1.88	0.074	-.2705375	5.443743
_Ilsdredre~1	.04917	.1744854	0.28	0.781	-.3136923	.4120323
_Ilsdredre~2	-2.029129	1.570542	-1.29	0.210	-5.29525	1.236992
isnightrllib	-.0791428	.0577926	-1.37	0.185	-.199329	.0410435
isnightifwor	3.952962	2.088877	1.89	0.072	-.391096	8.297021
ispas1rec	-.8806373	.6443105	-1.37	0.186	-2.220554	.4592797
isifsup	-.7353308	.835918	-0.88	0.389	-2.473717	1.003056
isiffam	.9064637	.6406229	1.41	0.172	-.4257846	2.238712
isage	-.4324577	.5493029	-0.79	0.440	-1.574796	.7098802
_Iisdred_2	.7456382	.3945228	1.89	0.073	-.0748169	1.566093
isexit	-3.856032	1.611587	-2.39	0.026	-7.207511	-.5045527
_cons	103.1957	97.99118	1.05	0.304	-100.5881	306.9795

Figure 7.1: Full Bayesian Meta-regression

BAYES, GIBBS SAMPLING, RANDOM MODEL, 19 YEAR OLDS (50,000 iterations, burn-in period of 1000 iterations), RANDOM SEED 1

NOTE THAT MD2 HAS BEEN DELETED FROM THE FILE BECAUSE IT WAS IDENTIFIED AS AN OUTLIER.

$$\text{logadjRR10}_{ij} \sim \mathrm{N}(XB, \Omega)$$

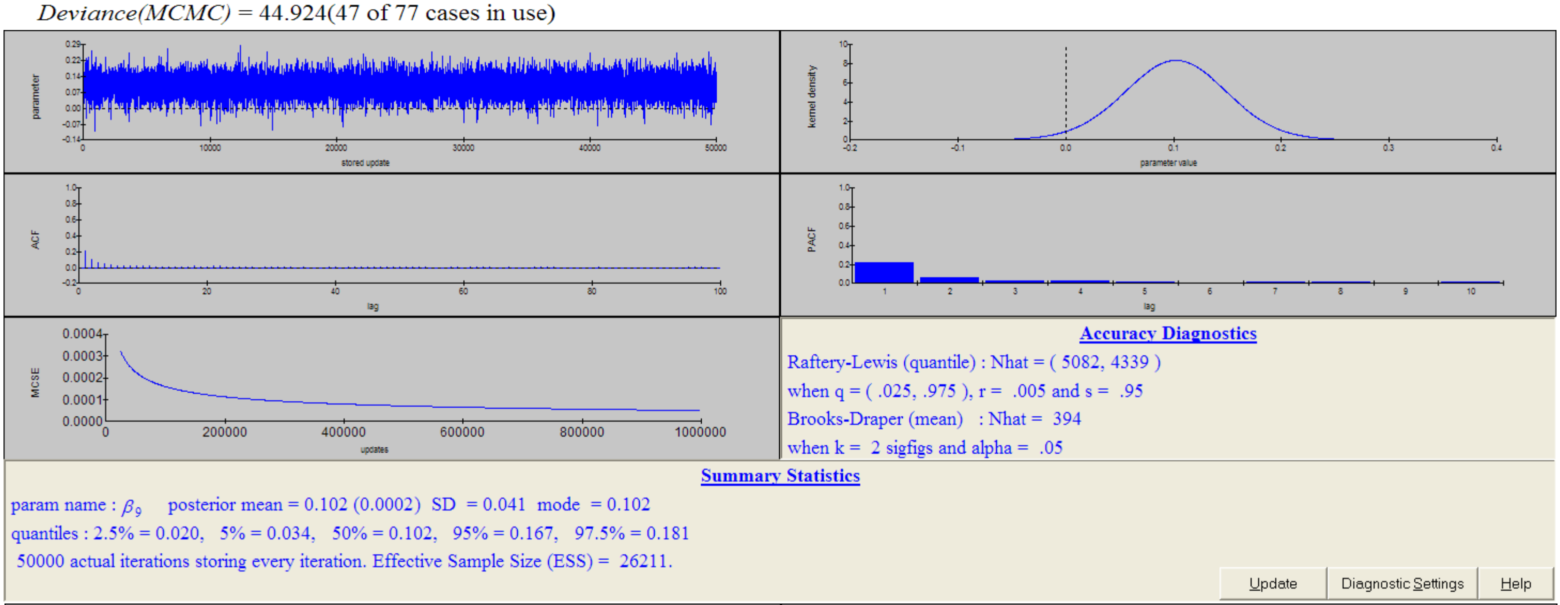
$$\begin{aligned}\text{logadjRR10}_{ij} = {} & \beta_{0j}\text{CONS} + -0.051(0.042)\text{year}_j + 0.176(0.213)\text{g}_j + -0.116(0.274)\text{m}_j + -0.032(0.055)\text{lslengthmin}_j + 0.037(0.042)\text{lslengthmax}_j + \\ & -0.005(0.006)\text{lspar1min}_j + 0.084(0.229)\text{lspar2rec\_1}_j + 0.102(0.041)\text{lsnightrl}_j + -0.102(0.371)\text{lsnightifsup\_1}_j + \\ & 0.265(0.492)\text{lspas1rec\_1}_j + 0.315(1.201)\text{lsiffam\_1}_j + -2.603(1.452)\text{lsifdrivered\_1}_j + -0.021(0.262)\text{lsage}_j + -0.060(0.210)\text{lsredu}_j + \\ & 2.543(1.085)\text{lsexit\_1}_j + 0.048(0.149)\text{lsdredrec\_1}_j + -2.072(1.241)\text{lsdredrec\_2}_j + -0.078(0.049)\text{isnightrllib}_j + \\ & 3.966(1.660)\text{isnightifwor\_1}_j + -0.890(0.540)\text{ispas1rec\_1}_j + -0.721(0.674)\text{isifsup\_1}_j + 0.924(0.535)\text{isiffam\_1}_j + \\ & -0.422(0.466)\text{isage}_j + 0.731(0.337)\text{isdred\_2}_j + -3.803(1.310)\text{isexit\_1}_j + e_{1ij}\text{logadjRR10se}_j\end{aligned}$$

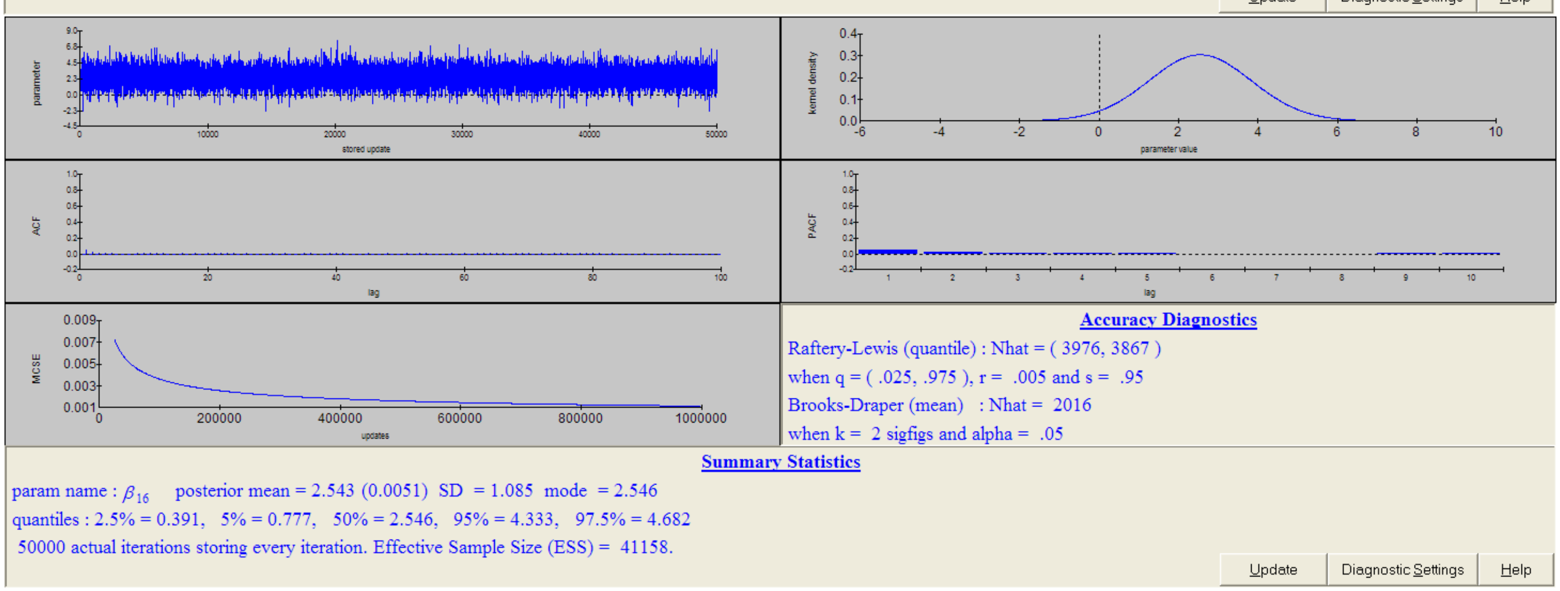
$$\beta_{0j} = 109.521(83.920) + u_{0j}$$

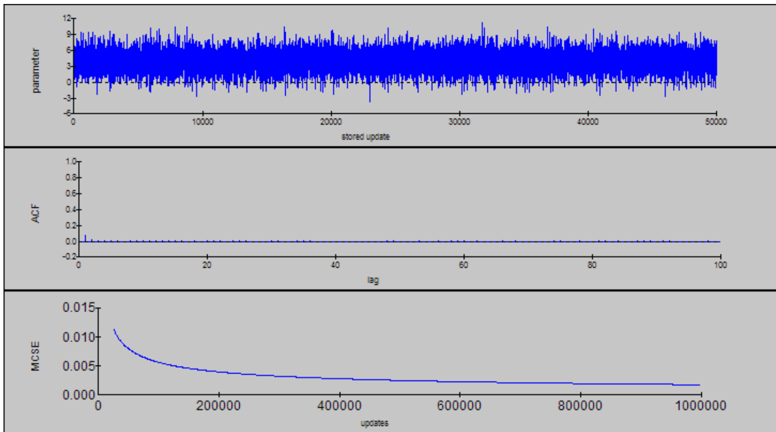
$$\left[u_{0j}\right] \sim \mathrm{N}(0,\ \Omega_u) \ :\ \Omega_u = \left[0.019(0.027)\right]$$

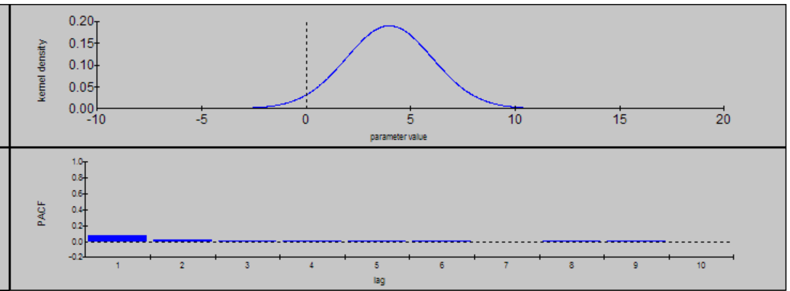
$$\left[e_{1ij}\right] \sim \mathrm{N}(0,\ \Omega_e) \ :\ \Omega_e = \left[0.586(0.241)\right]$$

Deviance(MCMC) = 44.924(47 of 77 cases in use)









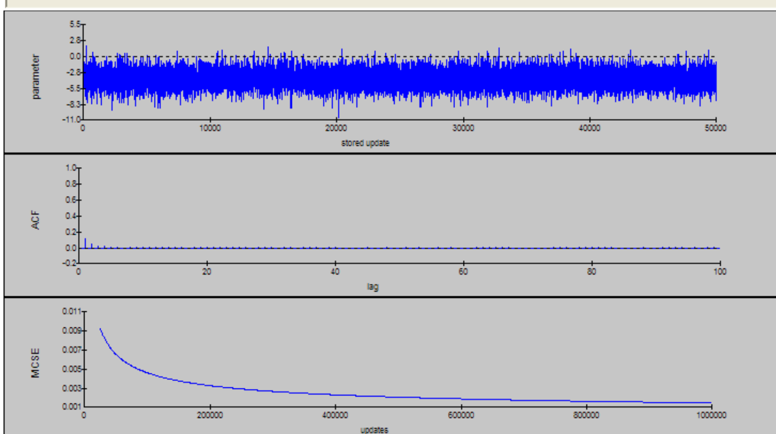
Accuracy Diagnostics

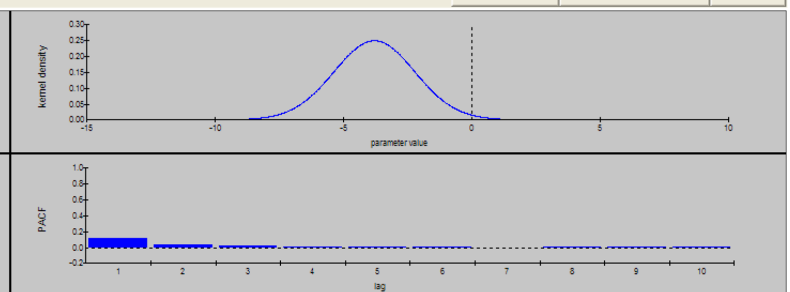
Raftery-Lewis (quantile) : Nhat = (3983, 3983)
when q = (.025, .975), r = .005 and s = .95
Brooks-Draper (mean) : Nhat = 4881
when k = 2 sigfigs and alpha = .05

Summary Statistics

param name : β_{20} posterior mean = 3.966 (0.008) SD = 1.660 mode = 3.972
quantiles : 2.5% = 0.675, 5% = 1.259, 50% = 3.963, 95% = 6.690, 97.5% = 7.237
50000 actual iterations storing every iteration. Effective Sample Size (ESS) = 40759.

Update | Diagnostic Settings | Help





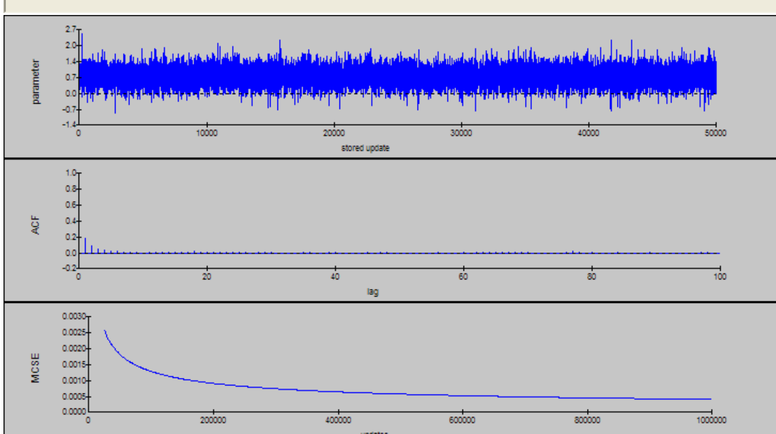
Accuracy Diagnostics

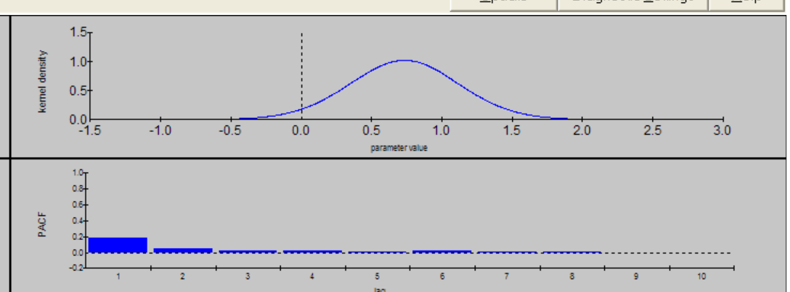
Raftery-Lewis (quantile) : Nhat = (4030, 4063)
when q = (.025, .975), r = .005 and s = .95
Brooks-Draper (mean) : Nhat = 3294
when k = 2 sigfigs and alpha = .05

Summary Statistics

param name : β_{26} posterior mean = -3.803 (0.0065) SD = 1.310 mode = -3.796
quantiles : 2.5% = -6.383, 5% = -5.950, 50% = -3.805, 95% = -1.674, 97.5% = -1.195
50000 actual iterations storing every iteration. Effective Sample Size (ESS) = 35255.

Update | Diagnostic Settings | Help





Accuracy Diagnostics

Raftery-Lewis (quantile) : Nhat = (4722, 4411)
when q = (.025, .975), r = .005 and s = .95
Brooks-Draper (mean) : Nhat = 25505
when k = 2 sigfigs and alpha = .05

Summary Statistics

param name : β_{25} posterior mean = 0.731 (0.0018) SD = 0.337 mode = 0.735
quantiles : 2.5% = 0.057, 5% = 0.179, 50% = 0.732, 95% = 1.283, 97.5% = 1.394
50000 actual iterations storing every iteration. Effective Sample Size (ESS) = 28610.

Update | Diagnostic Settings | Help

BAYES, GIBBS SAMPLING, RANDOM MODEL, 19 YEAR OLDS (50,000 iterations, burn-in period of 1000 iterations), RANDOM SEED 2

$$\text{logadjRR10}_{ij} \sim \text{N}(XB, \Omega)$$

$$\text{logadjRR10}_{ij} = \beta_{0j}\text{CONS} + -0.051(0.042)\text{year}_j + 0.178(0.211)\text{g}_j + -0.113(0.274)\text{m}_j + -0.031(0.055)\text{lslengthmin}_j + 0.037(0.041)\text{lslengthmax}_j + -0.005(0.006)\text{lspar1min}_j + 0.088(0.226)\text{lspar2rec\_1}_j + 0.102(0.040)\text{lsnightrl}_j + -0.100(0.368)\text{lsnightifsup\_1}_j + 0.262(0.487)\text{lspas1rec\_1}_j + 0.321(1.197)\text{lsiffam\_1}_j + -2.607(1.461)\text{lsifdrivered\_1}_j + -0.022(0.258)\text{lsage}_j + -0.061(0.209)\text{lsredu}_j + 2.544(1.076)\text{lsexit\_1}_j + 0.047(0.149)\text{lsdredrec\_1}_j + -2.072(1.233)\text{lsdredrec\_2}_j + -0.079(0.049)\text{isnightrllib}_j + 3.986(1.642)\text{isnightifwor\_1}_j + -0.894(0.532)\text{ispas1rec\_1}_j + -0.713(0.670)\text{isifsup\_1}_j + 0.925(0.527)\text{isiffam\_1}_j + -0.419(0.462)\text{isage}_j + 0.730(0.333)\text{isdred\_2}_j + -3.811(1.299)\text{isexit\_1}_j + e_{1ij}\text{logadjRR10se}_j$$

$$\beta_{0j} = 109.415(82.965) + u_{0j}$$

$$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.019(0.028)\right]$$

$$\left[e_{1ij}\right] \sim \text{N}(0,\ \Omega_e)\ :\ \Omega_e = \left[0.579(0.244)\right]$$

Deviance(MCMC) = 44.406(47 of 77 cases in use)

BAYES, GIBBS SAMPLING, RANDOM MODEL, 19 YEAR OLDS (50,000 iterations, burn-in period of 1000 iterations), RANDOM SEED 3

$$\text{logadjRR10}_{ij} \sim \text{N}(XB, \Omega)$$

$$\text{logadjRR10}_{ij} = \beta_{0j}\text{CONS} + -0.051(0.043)\text{year}_j + 0.177(0.213)\text{g}_j + -0.113(0.274)\text{m}_j + -0.031(0.055)\text{lslengthmin}_j + 0.037(0.042)\text{lslengthmax}_j + -0.005(0.006)\text{lspar1min}_j + 0.085(0.228)\text{lspar2rec\_1}_j + 0.102(0.040)\text{lsnightrl}_j + -0.100(0.369)\text{lsnightifsup\_1}_j + 0.263(0.490)\text{lspas1rec\_1}_j + 0.325(1.185)\text{lsiffam\_1}_j + -2.591(1.459)\text{lsifdrivered\_1}_j + -0.024(0.262)\text{lsage}_j + -0.060(0.210)\text{lsredu}_j + 2.547(1.083)\text{lsexit\_1}_j + 0.046(0.150)\text{lsdredrec\_1}_j + -2.078(1.237)\text{lsdredrec\_2}_j + -0.078(0.050)\text{isnightrllib}_j + 3.977(1.655)\text{isnightifwor\_1}_j + -0.892(0.537)\text{ispas1rec\_1}_j + -0.708(0.669)\text{isifsup\_1}_j + 0.925(0.533)\text{isiffam\_1}_j + -0.420(0.466)\text{isage}_j + 0.727(0.334)\text{isdred\_2}_j + -3.812(1.312)\text{isexit\_1}_j + e_{1ij}\text{logadjRR10se}_j$$

$$\beta_{0j} = 110.343(84.802) + u_{0j}$$

$$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.020(0.029)\right]$$

$$\left[e_{1ij}\right] \sim \text{N}(0,\ \Omega_e)\ :\ \Omega_e = \left[0.576(0.241)\right]$$

Deviance(MCMC) = 44.104(47 of 77 cases in use)

BAYES, GIBBS SAMPLING, RANDOM MODEL, 19 YEAR OLDS (50,000 iterations, burn-in period of 1000 iterations), RANDOM SEED 4

$$\text{logadjRR10}_{ij} \sim \mathrm{N}(XB, \Omega)$$

$$\begin{aligned}\text{logadjRR10}_{ij} = {} & \beta_{0j}\text{CONS} + -0.051(0.042)\text{year}_j + 0.175(0.214)\text{g}_j + -0.113(0.274)\text{m}_j + -0.032(0.055)\text{lslengthmin}_j + 0.037(0.042)\text{lslengthmax}_j + \\ & -0.005(0.006)\text{lspar1min}_j + 0.084(0.230)\text{lspar2rec\_1}_j + 0.101(0.040)\text{lsnightrl}_j + -0.101(0.372)\text{lsnightifsup\_1}_j + \\ & 0.260(0.494)\text{lspas1rec\_1}_j + 0.312(1.189)\text{lsiffam\_1}_j + -2.608(1.466)\text{lsifdrivered\_1}_j + -0.023(0.262)\text{lsage}_j + -0.059(0.210)\text{lsredu}_j + \\ & 2.547(1.076)\text{lsexit\_1}_j + 0.048(0.149)\text{lsdredrec\_1}_j + -2.084(1.240)\text{lsdredrec\_2}_j + -0.078(0.050)\text{isnightrllib}_j + \\ & 3.986(1.659)\text{isnightifwor\_1}_j + -0.892(0.534)\text{ispas1rec\_1}_j + -0.713(0.677)\text{isifsup\_1}_j + 0.925(0.531)\text{isiffam\_1}_j + \\ & -0.416(0.467)\text{isage}_j + 0.728(0.335)\text{isdred\_2}_j + -3.807(1.301)\text{isexit\_1}_j + e_{1ij}\text{logadjRR10se}_j\end{aligned}$$

$$\beta_{0j} = 110.044(84.443) + u_{0j}$$

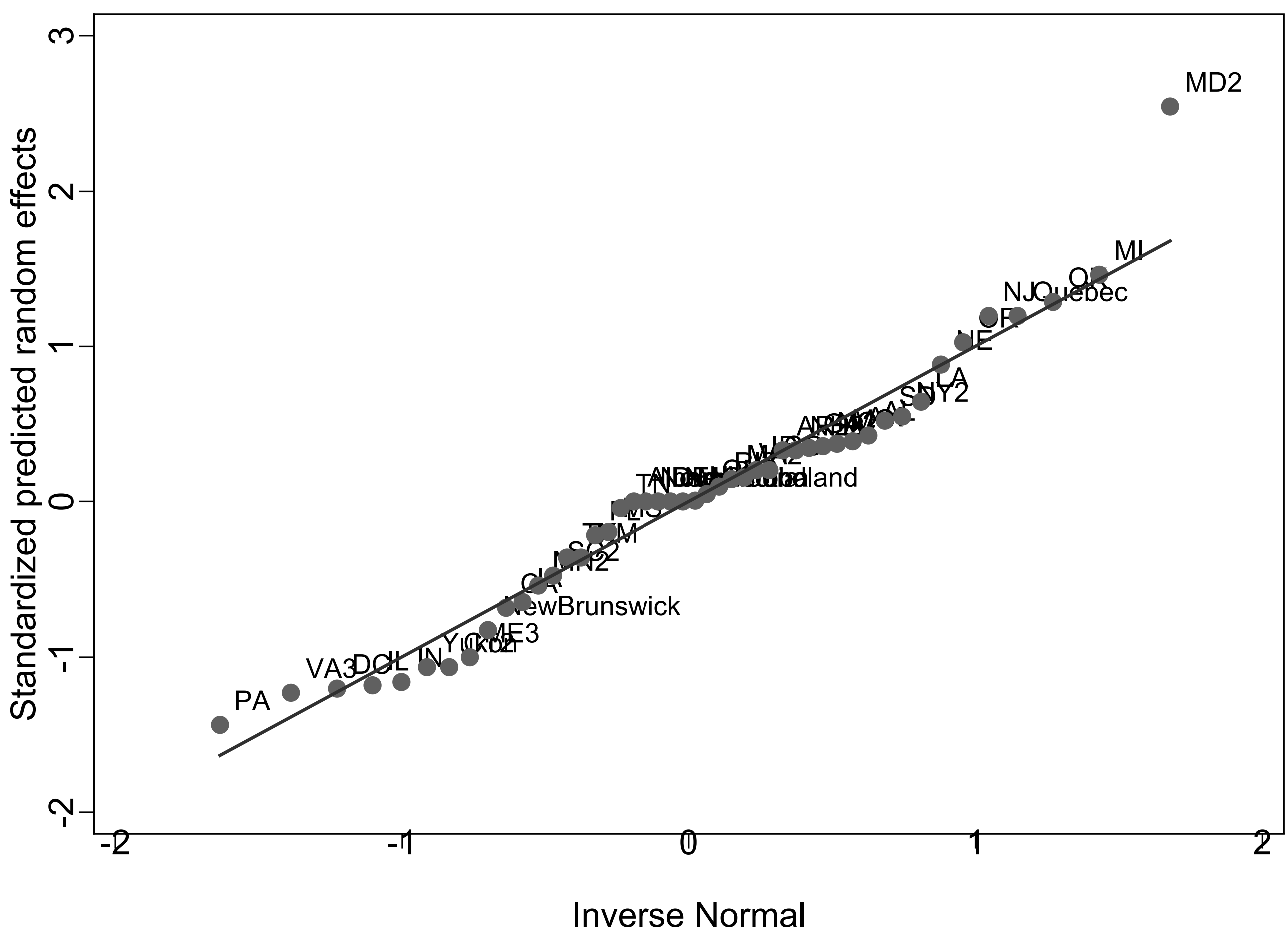
$$\left[u_{0j}\right] \sim \mathrm{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.020(0.027)\right]$$

$$\left[e_{1ij}\right] \sim \mathrm{N}(0,\ \Omega_e)\ :\ \Omega_e = \left[0.582(0.251)\right]$$

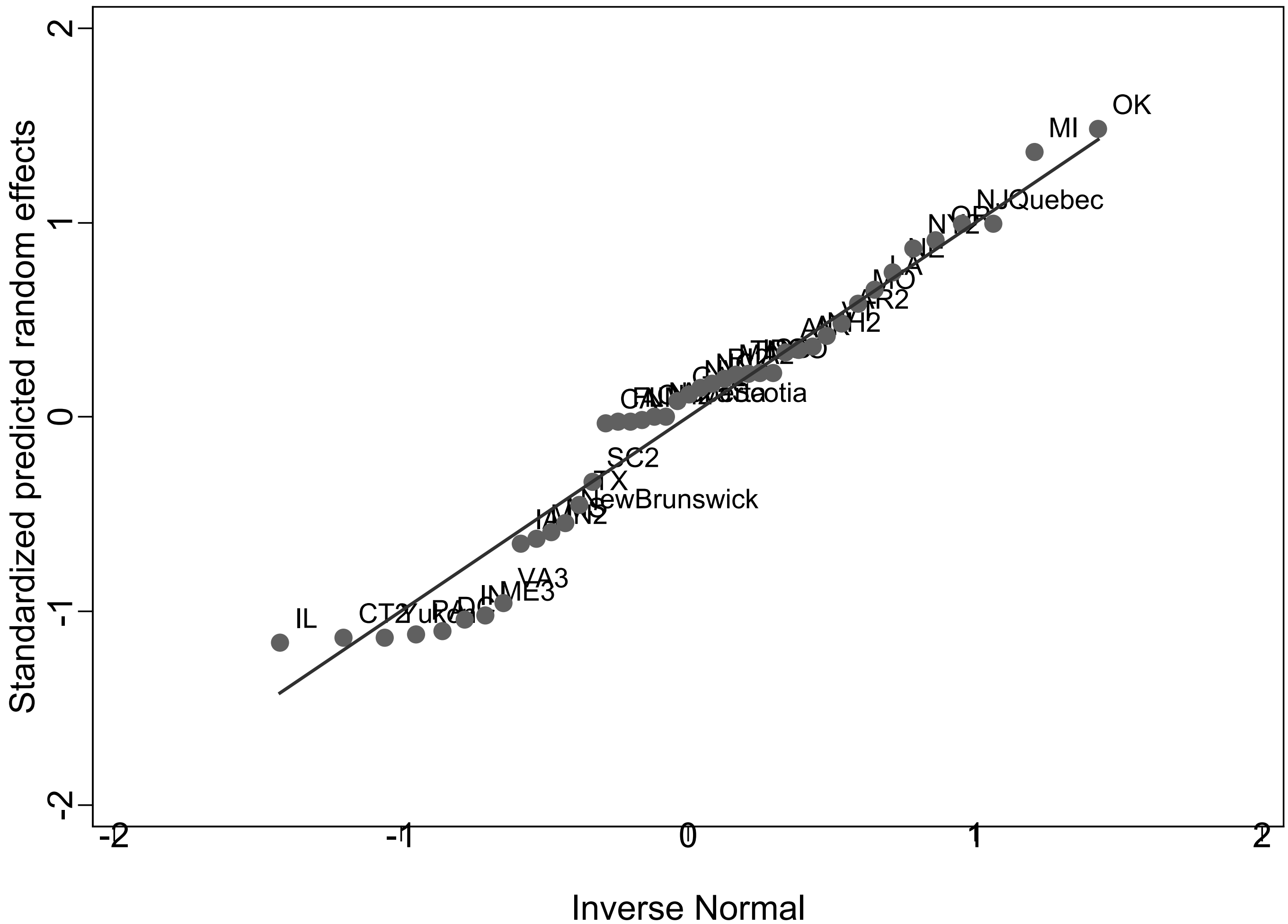
Deviance(MCMC) = 44.452(47 of 77 cases in use)

Figure 7.2: Normal probability plot of standardized shrunken residuals

With outlier MD2



Without outlier MD2



Appendix 8: Number of GDL components per program

Legend

juris – jurisdiction
date – GDL implementation date (year)
iihs – iihs rating: good($\geq$6 points) fair(4-5 points) marginal(2-3 points) poor(0-1 points)
gdl – gdl stages: both learner and intermediate stage

Learner Stage

- lsc1 - Learner stage entry age ($\geq$16)
- lsc2 - Learner stage length - minimum mandatory holding period ($\geq$6 months)
- lsc3 - Learner stage length - minimum mandatory holding period ($\leq$3 months)
- lsc4 - Minimum amount of supervised driving ($\geq$30 hours)
- lsc5 - Minimum amount of supervised driving ($\geq$50 hours)
- lsc6 - Mandatory hours of driving at night and/or inclement weather/before age 16 (yes/no)
- lsc7 - Learner stage night restriction (begins at 9 or 10 pm)
- lsc8 - Learner stage night restriction (begins after 10 pm)
- lsc9 - Learner stage passenger restriction ($\leq$1 passengers allowed)
- lsc10 - Learner stage passenger restriction ($\leq$2 passengers allowed)
- lsc11 - Learner stage exit test (yes/no)

Intermediate Stage

- isc1 - Intermediate stage entry age – difference between entry age for learner stage and intermediate stage is $\geq$12 months
- isc2 - Intermediate stage – most conservative night restriction (begins at 9 or 10 pm)
- isc3 - Intermediate stage – most conservative night restriction (begins after 10 pm)
- isc4 - Intermediate stage passenger restriction ($\leq$1 passengers allowed)
- isc5 - Intermediate stage passenger restriction ($\leq$2 passengers allowed)
- isc6 - Intermediate stage exit test (yes/no)
- total – total number of 'x's checked off

Table 8.1: Count of number of GDL components per GDL program

juris	date	iihs	gdl	lsc1	lsc2	lsc3	lsc4	lsc5	lsc6	lsc7	lsc8	lsc9	lsc10	lsc11	isc1	isc2	isc3	isc4	isc5	isc6	total
MD1	1992		x														x				2
MA1	1992		x	x													x				3
MN1	1992		x												x						2
NY1	1992		x	x												x	x				4
ON	1994	m		x	x	x					x			x						x	6
NS	1994	f	x	x	x	x						x	x	x			x			x	9
NB	1996	m	x	x	x	x						x	x	x							7
FL	1996	f	x		x	x	x	x	x						x		x				8
VA1	1996		x		x	x									x						4
KY1	1996	m		x	x	x					x										4
CT1	1997			x		x															2
MI	1997	f	x		x	x	x	x	x						x		x				8
QC	1997	m	x	x	x	x					x										5
GA1	1997		x		x	x									x		x				5
NC	1997	g	x		x	x					x				x	x	x	x	x		9
IL	1998	g	x		x	x	x	x	x						x	x	x	x	x		11
LA	1998	f	x		x	x									x		x				5
NH1	1998		x			x											x				3
CA	1998	g			x	x	x	x	x								x	x	x		8
OH1	1998		x																		1
SC1	1998		x			x										x	x				4
VA2	1998		x		x	x									x						4
ME1	1998		x			x	x		x						x						5
BC1	1998	m		x		x					x	x	x	x							6
MA2	1998	g	x	x	x	x	x				x						x	x	x		9
IN	1999	f	x			x									x		x	x	x		6
IA	1999	f	x		x	x			x						x		x				6
MN2	1999	g	x		x	x	x		x						x		x	x	x		9
NE	1999	g	x		x	x			x						x		x	x	x		8
NF	1999	f	x	x	x	x					x	x	x	x			x			x	10
OH2	1999	g	x		x	x	x	x	x								x				7
RI1	1999		x	x	x	x											x				5
SD	1999	m	x		x	x				x	x					x	x				7
OK	1999	g	x		x	x	x		x								x	x	x		8
CO	1999	g			x	x	x	x	x						x		x	x	x		9
DE	1999	g	x	x	x	x	x	x	x			x	x			x	x	x	x		13
MD2	1999	g	x		x	x	x	x	x								x	x	x		9
UT1	1999		x				x		x								x				4
AR1	1999				x	x									x						3
ND	1999	m			x	x					x										3
NH2	1999	f	x			x											x				3
PA	1999	g	x	x	x	x	x	x									x				7
PE	2000	f	x		x	x						x	x								5
NM	2000	m	x		x	x	x	x	x								x	x	x		9
OR	2000	g	x		x	x	x	x							x		x	x	x		9
DC	2000	g	x	x	x	x	x		x	x	x						x	x	x		11
MS	2000	m	x		x	x															3
VT	2000	f	x		x	x	x		x						x	x	x	x	x		9
WI	2000	g	x		x	x	x		x				x				x	x	x		9
ME2	2000		x			x	x		x						x			x	x		7
YU	2000	g	x		x	x	x	x	x		x	x	x	x	x		x				12
ID	2001	m	x		x	x	x	x	x							x	x	x	x		10
MO	2001	g	x		x	x	x		x						x		x	x	x		9
NJ	2001	g	x	x	x	x					x	x	x		x		x	x	x		11
WV	2001	f	x		x	x									x		x				5
NV	2001	g	x		x	x	x	x	x							x	x	x	x		10
TN	2001	g	x		x	x	x	x	x	x	x				x	x	x	x	x		13
UT2	2001		x				x		x								x	x	x		6
VA3	2001	g	x		x	x	x		x								x	x	x		8
WA	2001	g	x		x	x	x	x	x						x		x	x	x		10
GA2	2002	g	x		x	x	x		x						x			x	x		8
TX	2002	f	x		x	x									x		x	x	x		7
SC2	2002	m	x		x	x	x		x		x					x	x	x	x		10
AR2	2002	m			x	x									x			x	x		5
AL	2002	f			x	x									x		x				4
MB	2003	f	x	x	x	x								x					x		6
NH3	2003		x														x	x	x		4
AB	2003	f	x		x	x					x			x	x					x	7
UT3	2003	f	x				x		x								x	x	x		6
VA4	2003		x		x	x	x		x								x	x	x		8
RI2	2003	g	x	x	x	x	x	x	x								x				8
NY2	2003	g	x	x	x	x										x	x		x		7
ME3	2003	g	x		x	x	x		x						x		x	x	x		9
CT2	2003	g	x	x	x	x												x	x		6
BC2	2003		x	x	x	x					x	x	x	x					x	x	10
AK	2004	g			x	x	x		x						x		x	x	x		8
SA	2005		x		x	x								x						x	5
WY	2005	f	x				x	x	x						x		x	x	x		8