A propensity score approach to estimating child restraint effectiveness in preventing mortality

Michael R. Elliott[*],
Department of Biostatistics, University of Michigan School of Public Health, M4041, SPH II, 1420 Washington Heights, Ann Arbor, MI 48109, USA

Dennis R. Durbin, and
TraumaLink Injury Research Center, The Children's Hospital of Philadelphia, Division of Emergency Medicine, Department of Pediatrics Center for Clinical Epidemiology and Biostatistics University of Pennsylvania

Flaura K. Winston
TraumaLink Injury Research Center The Children's Hospital of Philadelphia, Division of General Pediatrics, Department of Pediatrics Leonard Davis Institute for Health Economics University of Pennsylvania

Abstract

Confounding between the child's restraint use and driver behavior can bias restraint effectiveness estimates away from the null if survivable crashes are more common in certain restraint types. Analyzing only fatal crashes may introduce selection bias toward the null because any protective effects of a restraint type will underrepresent children in that restraint. A marginal-structural-model-type estimator suggests a 17% reduction in fatality risk for children aged 2 through 6 in child restraint systems relative to seat belts. This reduction is estimated at 22% when severe misuse of the restraint is excluded.

Keywords and Phrases

Marginal structural model; Selection bias; Confounding; Fatality; Child safety seat; Injury epidemiology

1. INTRODUCTION

Vehicle safety policy is largely driven by estimates of relative effectiveness among options for protection. For example, regulatory requirements for airbags (Federal Motor Vehicle Safety Standard 208, Occupant Crash Protection [49 CFR 571.208]) were supported by estimates of the supplemental protection afforded by airbags over seat belts alone. Similarly, state laws requiring the use of child restraints for children have relied on evidence (Arborgast et al. 2004, Durbin et al. 2003) demonstrating that child restraint effectiveness was greater than that of seat belts in protecting children in crashes. Arbogast et al. showed

[*]Corresponding author. Institute for Social Research, University of Michigan, mrelliot@umich.edu.

that children 12–47 months of age had a 78% reduction in injury risk when seated in forward-facing child restraints versus seat belts; Durbin et al. found that children aged 4 through 7 years had a 59% reduction in injury risk when seated in belt-positioning booster seats versus seat belts. Restraint effectiveness has often been described in terms of mortality reduction, but conflicting conclusions can result based on the analytical methods chosen for effectiveness estimation. For example, in a previous analysis of the effectiveness of child restraint seats (CRSs) relative to seat belts, Levitt (2005) used FARS data from 1975 to 2003 and, by various methods, directly compared the mortality rates for child restraints and for seat belts for children ages 2 to 6 and could not demonstrate a difference in effectiveness relative to no restraint. This analysis received considerable attention in the popular press. In a New York Times article Dubner and Levitt (2005) declared that money spent on child restraint systems would be better spent on back-seat DVD players to force children to sit still in the back seat. ABC News' Prime Time (http://a.abcnews.com/Primetime/Story?id=1842987&page=1) and PBS's Travis Smiley show (http://www.pbs.org/kcet/tavissmiley/archive/200508/20050802_levitt.html) also aired stories promoting Levitt's research. However, the studies of Durbin et al. (2003) and Levitt (2005) may have such startlingly different conclusions because of differing study methodologies. No study to date has compared estimation methods in order to assess which provides the least biased estimate. We will explore the issue of bias in estimating the relative effectiveness of child restraint systems over seat belt restraint for young children.

Child restraint systems (CRS), as distinct from seat belts, include child safety seats and boosters seats and are designed to address the biomechanical and size safety needs of children as seat belt fit is poor for children under 4 feet 9 inches tall (nearly all children under age 8 years). Unfortunately, the derivation of restraint effectiveness estimates based on laboratory testing is limited by the inadequate biofidelity of the anthropomorphic dummy used in testing, by the lack of measurement of or inaccurate measurement of injury risk in the dummy, and by the relative simplicity of the laboratory test configuration as compared with real-world crashes and restraint system use. Consequently, comparisons of effectiveness between child restraints and seat belts are largely relegated to statistical analysis of real world crash databases.

Estimating effectiveness of child restraint systems through analysis of crash databases is problematic due to the association between how passengers are restrained in a given crash and whether that crash will be in the given database. The primary sources of US-level data available for assessing mortality associated with restraint type include the Fatality Analysis Reporting System (FARS) (National Highway Traffic Safety Administration, 2005), the National Automotive Sampling System General Estimates System (NASS GES) (National Highway Traffic Safety Administration, 2000), and the National Automotive Sampling System Crashworthiness Data System (NASS CDS) (National Highway Traffic Safety Administration, 1997). FARS is a census of vehicular crashes in the US in which at least one person died (not necessarily the child passenger) within the 50 States, the District of Columbia, and Puerto Rico. FARS has a sufficient number of outcomes of fatal child injuries for analysis but has a biased selection of crashes in that inclusion of crashes is associated with the outcome of interest, mortality. NASS GES and NASS CDS compile data from a nationally-representative sample of police-reported crashes (restricted to crashes in which at

least one vehicle was non-drivable in the case of NASS CDS). While both FARS and NASS
GES rely on police reports as the primary source of data, NASS CDS includes data from
detailed crash investigations by trained investigators supplemented by review of medical
records. Thus the methodology of NASS CDS as compared to NASS GES results in more
detailed and reliable data regarding restraint status and crash circumstances and, therefore, is
the more scientifically rigorous of the two regarding restraint. However, while NASS CDS
contains information on children in fatal crashes, it is a probability sample with a relatively
small sampling fraction. Despite oversampling of more severe crashes, only about 1–3% of
fatal crashes involving children are included in NASS CDS, compared with 100% in FARS.
Including FARS adds enormously to power to detect effects of restraint type on risk of
death, since a substantial fraction of FARS crashes belong to the relevant set of potentially-
fatal crashes. In the following, we will demonstrate the biases associated with use of these
databases individually in deriving effectiveness estimates and will present a more robust
estimation procedure that depends upon the use of both databases.

1.1 Selection bias and confounding in restraint effectiveness estimation

To understand the nature of the problem of bias introduced by database selection, we turn to
the potential outcomes paradigm (Rubin, 1974; Rubin 1978). Following the concepts of the
Rubin Causal Model (Holland, 1988), we want to compare the risk of death for a child
restrained in a CRS with the risk of death for a child restrained in a belt *among the subset of
crashes which would have resulted in death in at least one of the restraint types*. Failure to
condition on these "potentially fatal" crashes means that any association between CRS use
and fatality may be confounded with driver behavior. That is, a population-based
denominator for the CRS users to estimate CRS risk will contain a disproportionate number
of crashes in which the child would have survived in either restraint type. A natural
alternative is to restrict analyses to crashes in which fatalities occurred, since they clearly are
all "potentially fatal" crashes. However, using data from crashes in which at least one fatality
occurred will remove from the CRS denominator all crashes for which the CRS would have
been protective, unless someone else died in the crash.

Table 1 illustrates the issue, under the simplifying assumption that CRS are never harmful.
Let Y represent whether or not a restrained child dies(=1) or survives(=0) in a crash, and T
represent whether or not the child is restrained in a CRS(=1) or a seat belt(=0). We also
define W to be an indicator of whether there was a fatality other than the index child in the
crash, either in the child's vehicle or another vehicle in the crash. Next, let $Y(T)$ represent
the two potential outcomes that a child would have had in the crash: $Y(1)$ indicates whether
or not s/he would have died in the crash had s/he been in a CRS, and $Y(0)$ indicates whether
or not s/he would have died in the crash had s/he been in a belt. We thus classify six types of
crashes: "always fatal," "fatal in belt only," and "always survivable," by whether or not
someone in another vehicle dies. If crash type is independent of restraint type use, i.e., if
crash severity is "randomized" with respect to restraint type (column A in Table 1), then a
population-based cohort analysis using randomly sampled crashes from the entire population
provides consistent estimates of the true protective effect for the CRS. A FARS-only
analysis will typically be approximately unbiased as well, although a slight bias toward the
null will occur to the degree that child-only deaths are common relative to child fatalities

where at least one other person in the crash dies. However, we cannot assume that restraint use is randomized. If crashes in which someone else dies beside the child are less common in CRS-restrained children (column B in Table 1), a FARS-only analysis will be biased toward the null, whereas a population-based cohort analysis will be biased away from the null. If the association between restraint use and crash severity is reversed, (column C in Table 1), the direction of the bias switches in a FARS-only and a population-based cohort analysis. We anticipate that children restrained in CRS are less likely to be in more severe crashes (the Table 1, column B scenario), suggesting that a FARS-only analysis might underestimate the effectiveness of CRS relative to belts, whereas a standard cohort analysis might overestimate the effectiveness.

To overcome these complementary problems, we combine 1998–2003 data from both the FARS and the National Automotive Sampling System (NASS) to obtain a full population cohort of children in towaway crashes. We then propose a method using a propensity score to obtain estimates of restraint effectiveness that should be protective against both sources of bias. Typically the covariates used to estimate the propensity scores can be included in a regression model to yield similar adjusted estimates; however, by treating the potential outcomes cells as coarsened cells in a contingency table stratified by crash fatality status, we obtain a simple estimator of CRS effectiveness that incorporates data from non-fatal crashes but is more robust than a standard adjusted relative risk estimator that pools across crash fatality status. We compare our method with a standard cohort analysis using both the FARS/NASS data combined and a FARS-only analysis; for the latter we also consider alternative methods proposed by Evans (1986) and Levitt and Porter (2001) to reduce selection bias in a FARS-only analysis. Because all state laws require child restraints for children under age 2, but some states still allow seat belt restraints for children over 2, this study was limited to children between the ages of 2 and 6. This population is also the same for which controversy has developed regarding CRS effectiveness.

2. A ROBUST CAUSAL ESTIMATOR OF RESTRAINT EFFECTIVENESS

In this section, we describe a robust causal estimator of restraint effectiveness. We use the term "causal" to denote that, under the assumption of "no unobserved confounders," it is a consistent estimator of the relative risk of death for a child in a CRS versus a child in a seat belt if restraint use type were assigned randomly to children in the population. Our method uses propensity scores to combine adjusted analyses across the strata of otherwise fatal and otherwise non-fatal crashes without having to explicitly model the adjustment covariates, similar to the recently popularized marginal structural models (Joffe et al. 2004; Robins, 1999; Robins et al. 2000) which use propensity to comply to treatment to create artificial populations of subjects randomized to different treatment arms in order to make regression-adjusted estimates of treatment effects without directly adjusting for confounders with treatment assignment in the mean regression model. An advantage of the propensity score approach is that it does not rely on the linearity assumption of a standard linear or logistic regression model; hence it allows for more flexible constructions. In particular, we can model propensities separately within fatality strata, allowing for differing "assignment mechanisms" in fatal vs. non-fatal crashes, and then average the results across the strata for a population-level effect. In addition, subjects at a given exposure level with extremely low or

extremely high probabilities of assignment for whom there are no comparable subjects with different exposure levels are dropped from the analysis, since we only want to compare subjects who have a chance of having either exposure (Leon et al. 2001).

Injury epidemiology would likely benefit from increased use of propensity score methodology. Injuries by their nature are sporadic and difficult to predict, and usually rare in most populations. Randomized trials with sufficient power to detect treatment effects are thus very expensive to mount. Methods which better accommodate the limitations of observational data are thus of great value in injury epidemiology.

2.1 Relative risk estimation with potential outcomes

All restrained children in a crash could be assigned to one of 4 outcomes: "unsurvivable" crashes in which they would die regardless of whether they were in a belt or a CRS ($Y(1) = Y(0) = 1$), "CRS-survivable" crashes in which they would only die if they were in a belt ($Y(1) = 0$, $Y(0) = 1$), "belt-survivable" crashes in which they would die only if they were in a CRS ($Y(1) = 1$, $Y(0) = 0$), and "survivable" crashes in which they would not die in either restraint type ($Y(1) = Y(0) = 0$). Denote $P(Y(1) = j, Y(0) = k) = \pi_{jk}$, $\Sigma_j \Sigma_k \pi_{jk} = 1$; a "potentially fatal" crash implies $Y(1) + Y(0) \geq 1$, $P(Y(1) + Y(0) \geq 1) = 1 - \pi_{00}$. If we could observe the joint potential outcomes for each subject, we would estimate the relative risk of death in a CRS versus a belt by

$$RR_1 = \frac{P(Y(1) = 1 | Y(1) + Y(0) \geq 1)}{P(Y(0) = 1 | Y(1) + Y(0) \geq 1)} = \frac{(\pi_{11} + \pi_{10})/(1 - \pi_{00})}{(\pi_{11} + \pi_{01})/(1 - \pi_{00})} = \frac{\pi_{11} + \pi_{10}}{\pi_{11} + \pi_{01}}. \quad (1)$$

2.2 Propensity scores

Of course, it is unreasonable to assume that restraint assignment is fully randomized. A more reasonable assumption is "no unobserved confounders": that, conditional on covariates **X**, such an assignment is random; i.e., independent of ($Y(1)$, $Y(0)$):

$$P(Y(1), Y(0), T = 1 | \mathbf{X}) = P(Y(1), Y(0) | X)P(T = 1 | \mathbf{X}) \quad (2)$$

This is the "balancing property": if (2) holds, then Rosenbaum and Rubin (1983) show that

$$P(Y(1), Y(0), T = 1 | Z(\mathbf{X})) = P(Y(1), Y(0) | Z(\mathbf{X}))P(T = 1 | Z(\mathbf{X})) \quad (3)$$

where $Z(\mathbf{X}) = P(T = 1 | \mathbf{X})$, the probability that a child is restrained in type T given covariates X. Once these propensity scores have been estimated, the data are stratified into propensity percentiles (typically quintiles), denoted by Z. Letting PF be an indicator for the condition $Y(1) + Y(0) \geq 1$ (i.e., being in a potentially fatal crash), we then have

$$RR_{1z} = \frac{P(Y(1) = 1 | PF = 1, Z = z)}{P(Y(0) = 1 | PF = 1, Z = z)} \tag{4}$$

$$= \frac{P(Y(1) = 1 | PF = 1, Z = z, T = 1)}{P(Y(0) = 1 | PF = 1, Z = z, T = 0)}$$

$$= \frac{P(Y(1) = 1 | Z = z, T = 1)/[1 - P(PF = 0 | Z = z, T = 1)]}{P(Y(0) = 1 | Z = z, T = 0)/[1 - P(PF = 0 | Z = z, T = 0)]}$$

$$= \frac{P(Y^{obs} = 1 | Z = z, T = 1)/[1 - P(PF = 0 | Z = z)]}{P(Y^{obs} = 1 | Z = z, T = 0)/[1 - P(PF = 0 | Z = z)]}$$

$$= \frac{P(Y^{obs} = 1 | Z = z, T = 1)}{P(Y^{obs} = 1 | Z = z, T = 0)}$$

where the first equality defines our causal estimator (1) stratified by propensity score, the second equality follows from the balancing property (2) of the propensity score, the third equality from the definition of a conditional distribution, and the fourth equality from the definition of our potential outcome. Thus, within the strata defined by Z, the observed death rates within each restraint type are consistent estimators of the numerators and denominator of (1). An overall estimate of relative risk can then be obtained as

$$RR_1 = \sum_z RR_{1z} P(Z = z)$$

which is the mean of these relative risks, $\sum_z RR_{1z}/5$ if the quintile cutoffs are exact. Alternatively, the values of RR_{1z} can be considered to determine if there is effect modification with respect to restraint use. It is important to note at this point that the relative risk estimate obtained in this manner will be asymptotically equivalent to those obtained under a standard multivariate Poisson regression model adjusted for covariates X if the linear model is correct; Section 2.4 develops an alternative "robust" estimator that assumes randomization holds only within the "other fatality" (W) strata.

The propensity or "balancing" score approach depends on correct modeling of $Z(\mathbf{X}) = P(T = 1 | \mathbf{X})$. In our manuscript we assume a logistic regression model:

$$P(T = 1 | \mathbf{X}) = \frac{\exp(\alpha + \mathbf{X}'\beta)}{1 + \exp(\alpha + \mathbf{X}'\beta)} \tag{5}$$

where α and β are estimated from the data and the covariates used in $\mathbf{X}$ are chosen by stepwise regression. If the model is approximately correct, the stratified relative risk measures should differ little after adjustment for $\mathbf{X}$. It is important that only "pre-treatment" covariates be used in estimating the propensity score, in order to avoid absorbing the treatment effect into factors that are on the causal pathway to the outcome; hence factors such as crash severity should not be included in $\mathbf{X}$.

2.3 A robust relative risk estimator

A more robust alternative to (4) assumes that the balancing property (2) of the propensity score Z only holds within fatality strata W:

$$P(Y(1), Y(0), T = 1 | X, W) = P(Y(1), Y(0) | X, W) P(T = 1 | X, W)$$

This is akin to columns (B) and (C) in Table 1, where the distribution of potential outcomes is equal only within fatality strata, not across them as in column (A). Under this constraint, the relative risk estimator is given by:

$$RR_2 = \frac{P(Y(1) = 1 | PF = 1)}{P(Y(0) = 1 | PF = 1)} = \sum_z \frac{\sum_{w=0}^{1} P\left(Y^{obs} = 1 \middle| W = w, Z = z, T = 1\right) P(W = w | Z = z)}{\sum_{w=0}^{1} P\left(Y^{obs} = 1 \middle| W = w, Z = z, T = 0\right) P(W = w | Z = z)} P(Z = z).$$

The derivation of RR_2 is as follows. From the law of total probability we have:

$$RR_2 = \frac{P(Y(1) = 1 | PF = 1)}{P(Y(0) = 1 | PF = 1)}$$

$$= \frac{[P(Y(1) = 1 | PF = 1, W = 1) P(W = 1 | PF = 1) + P(Y(1) = 1 | PF = 1, W = 0) P(W = 0 | PF = 1)]}{[P(Y(0) = 1 | PF = 1, W = 1) P(W = 1 | PF = 1) + P(Y(0) = 1 | PF = 1, W = 0) P(W = 0 | PF = 1)]}$$

$$= \sum_z \frac{[P(Y(1) = 1 | PF = 1, W = 1, Z = z) P(W = 1 | PF = 1, Z = z) + P(Y(1) = 1 | PF = 1, W = 0, Z = z) P(W = 0 | PF = 1, Z = z)]}{[P(Y(0) = 1 | PF = 1, W = 1, Z = z) P(W = 1 | PF = 1, Z = z) + P(Y(0) = 1 | PF = 1, W = 0, Z = z) P(W = 0 | PF = 1, Z = z)]}$$

$$\times P(Z = z$$

$$)$$

From balancing property of the propensity score within stratum W we have:

$$\sum_z \frac{[P(Y(1) = 1 | PF = 1, W = 1, Z = z) P(W = 1 | PF = 1, Z = z) + P(Y(1) = 1 | PF = 1, W = 0, Z = z) P(W = 0 | PF = 1, Z = z)]}{[P(Y(0) = 1 | PF = 1, W = 1, Z = z) P(W = 1 | PF = 1, Z = z) + P(Y(0) = 1 | PF = 1, W = 0, Z = z) P(W = 0 | PF = 1, Z = z)]} P$$

$$(Z = z)$$

$$= \sum_z \frac{[P(Y(1) = 1 | PF = 1, W = 1, Z = z, T = 1) P(W = 1 | PF = 1, Z = z) + P(Y(1) = 1 | PF = 1, W = 0, Z = z, T = 1) P(W = 0 | PF = 1, Z = z)]}{[P(Y(0) = 1 | PF = 1, W = 1, Z = z, T = 0) P(W = 1 | PF = 1, Z = z) + P(Y(0) = 1 | PF = 1, W = 0, Z = z, T = 0) P(W = 0 | PF = 1, Z = z)]}$$

$$\times P(Z = z$$

$$)$$

Finally, from the definition of the potential outcome and Bayes' Theorem we have

$$\sum_z \frac{[P(Y(1)=1|PF=1,W=1,Z=z,T=1)P(W=1|PF=1,Z=z) + P(Y(1)=1|PF=1,W=0,Z=z,T=1)P(W=0|PF=1,Z=z)]}{[P(Y(0)=1|PF=1,W=1,Z=z,T=0)P(W=1|PF=1,Z=z) + P(Y(0)=1|PF=1,W=0,Z=z,T=0)P(W=0|PF=1,Z=z)]} P$$

$(Z = z)$

$$= \sum_z \frac{\left[\dfrac{P\left(Y^{obs}=1 \middle| W=1,Z=z,T=1\right)}{P(PF=1|W=1,Z=z)} P(PF=1|W=1,Z=z) \dfrac{P(W=1|Z=z)}{P(PF=1|Z=z)} + \dfrac{P\left(Y^{obs}=1,W=0,Z=z,T=1\right)}{P(PF=1|W=0,Z=z)} P(PF=1|W=0,Z=z) \dfrac{P(W=0|Z=z)}{P(PF=1|Z=z)} \right]}{\left[\dfrac{P\left(Y^{obs}=1 \middle| W=1,Z=z,T=0\right)}{P(PF=1|W=1,Z=z)} P(PF=1|W=1,Z=z) \dfrac{P(W=1|Z=z)}{P(PF=1|Z=z)} + \dfrac{P\left(Y^{obs}=1,W=0,Z=z,T=0\right)}{P(PF=1|W=0,Z=z)} P(PF=1|W=0,Z=z) \dfrac{P(W=0|Z=z)}{P(PF=1|Z=z)} \right]} P$$

$(Z = z)$

$$\sum_z \frac{\sum_{w=0}^1 P\left(Y^{obs}=1 \middle| W=w,Z=Z,T=1\right) P(W=w|Z=z)}{\sum_{w=0}^1 P\left(Y^{obs}=1 \middle| W=w,Z=Z,T=0\right) P(W=w|Z=z)} P(Z=z$$

$)$

where the third equality follows from algebraic cancellations.

2.4 Estimation and inference

Estimators utilize case weights to reflect unequal probabilities of selection in the NASS-CDS dataset; case weights for the FARS cases are set to 1 to reflect their certainty sampling. Confidence intervals for the relative risk estimates are obtained via a bootstrapping procedure. Resampling is done at the cluster (crash) level, within each of the propensity score strata: this accommodates 1) the case weights in the NASS sample, 2) the clustering of the FARS sample by crash and the NASS sample by primary sampling unit, and 3) the need to treat the propensity scores as ancillary statistics (Rubin, 1979; Rubin and Thomas 1996). (The FARS data are a census of all crashes, and thus do not have sampling variability from a finite population sampling perspective. However, we consider the FARS crashes drawn from a hypothetical infinite superpopulation of fatal crashes, and thus resample both FARS and NASS crashes for inference.)

2.5 Alternative methods

We compare our results with three existing methods that have been used to estimate the relative effectiveness of restraints in reducing risk of fatality in passenger vehicle crashes: a cohort analysis, the restricted sample method of Levitt and Porter (2001), and a matched case-control analysis. We created a complete cohort sample of children in the US who were in towaway crashes by combining NASS dataset and the FARS dataset (Elliott et al. 2006). To reduce confounding between "potentially fatal" crashes and observed restraint use, our cohort analysis adjusted for child age, driver age, seat row, vehicle type, and vehicle model year.

We reproduced the "restricted sample" method of Levitt and Porter, a method proposed by them to eliminate selection bias in the FARS data by restricting the analytic set to the subset of FARS crashes which are a) two-vehicle crashes where b) someone in the other vehicle died. This restriction relies on two assumptions: 1) the potential outcomes and safety device usage (restraint usage) are independent conditional on observed covariates, equivalent to the "conditional randomization" assumption in propensity score analysis, and 2) the survival status of subjects in other vehicles is independent of safety device usage of children, again conditional on the potential outcome and observed covariates. This second assumption may fail if drivers of vehicles who cause fatalities in other vehicles are also less likely to restrain young children correctly, which is a plausible scenario.

Finally, we conducted a matched case-control or conditional logistic regression analysis, which utilizes the subset of restrained children in which two or more children were present and at least one child died and one child survived. In this setting, the matched case-control analysis treats the vehicle-level risk of death as a nuisance parameter and computes a semiparametric likelihood that effectively conditions on the crash circumstances and avoids the need to make the assumptions underlying the Levitt and Porter approach. While in principle it allows adjustment for potential confounding by other factors that may systematically differ between children restrained in belts and children restrained in CRSs, the two most important factors – age and seat row – are almost completely confounded with CRS use in crashes in which multiple young children are present. Hence our matched case-control results are unadjusted and are similar to those obtained through "double sampling" (Evans 1986).

The full cohort analysis was conducted using case weights equal to the inverse of the probability of selection and adjusted to known crash totals to account for the oversampling of severe crashes in NASS-CDS. (Case weights in FARS were set to 1 consistent with the fact that the FARS is a census of all fatalities.) To adjust inference to account for the disproportional probability of selection of subjects and stratification and clustering of subjects by geographic region and vehicle, Taylor Series linearization estimates of the logistic regression parameter variances were calculated. For the FARS only cohort and Levitt and Porter analyses, generalized estimating equations were used to account for the clustering of subjects by vehicle. For the matched case-control analysis, Cox semiparametric regression models were used to accommodate the $m : n$ matching of cases to controls.

3. DATA SOURCES

A full population cohort is obtained by combining data from both the FARS and NASS Crashworthiness Data System (CDS) database. In order to be comparable with the NASS-CDS database described below, the 8% of restrained children in vehicles involved in fatal crashes that were still drivable were excluded from the analysis; further, to focus on the effectiveness of current restraint systems, only crashes between 1998 and 2003 were analyzed. A small number of crashes in which only non-occupants (e.g., pedestrians) died were also excluded. Within FARS, we identified 7,816 children aged 2–6 who were vehicle occupants restrained in a CRS or a seat belt in a non-drivable (towaway) passenger car, van, pickup truck, or sport utility vehicle that was involved in a crash with at least one passenger fatality between 1998 and 2003. Of these 7,816 children involved in fatal crashes, 1,096 (14%) were themselves fatalities. Approximately 5,000 vehicles per year are sampled as part of the NASS-CDS. Within NASS-CDS, we identified 1,436 children aged 2–6 who were restrained in a CRS or seat belt in a passenger car, van, pickup truck, or sport utility vehicle involved in a non-fatal crash sampled between 1998 and 2003. Because of the complex sample design of NASS-CDS, these 1,436 children represent 959,483 children meeting our inclusion criteria. Table 2 provides details about the samples used for the Levitt and Porter restricted sample and matched case-control analyses.

4. RESULTS

Table 3 provides descriptive statistics for the variables used in the analysis, overall and by the source of the data (FARS vs. NASS-CDS). The vast majority of children in towaway crashes (over 99%) survive; thus the distribution of NASS-CDS cases closely parallels that of the entire population.

The pre-crash covariates available in both datasets to construct a propensity score were child age, vehicle type (passenger car, pickup truck, van, sports utility vehicle), seat row, age of driver, gender of driver, and vehicle model year. A preliminary stepwise regression step showed that the only pre-crash covariates independently and significantly associated with CRS use were age of child and vehicle type (passenger car, pickup truck, van, sports utility vehicle). Hence these factors were used to generate the propensity score estimates. Table 4 shows that balance across the covariates within each propensity score quintile was largely achieved (the exception – a low rate of car seat usage among those front-seated in the second PS quintile – is largely due to several very low case-weight NASS cases in that cell). Figure 1 shows the distribution of propensity scores for CRS use by restraint type. Both restraint types span the range of propensity scores, allowing estimation of restraint effects within each propensity score quintile.

Table 5 shows the standard cohort relative risk estimators, unadjusted and adjusted, along with the causal relative risk estimator developed above, and relative risk estimator using the Levitt and Porter and double sampling methods. The standard cohort relative risk estimators are computed using the FARS data only and the full population (FARS and NASS-CDS combined); the robust causal estimator requires the full population data; and the Levitt and Porter and matched case-control estimators use only FARS data. When severe misuse of

restraints are included, none of the estimates show a statistically significant difference at the $a = .05$ level, although the adjusted standard cohort relative risk estimator approaches significance (RR = 0.69, 95% CI = 0.46, 1.02). Removing the small fraction of FARS seat belt users who only used a shoulder belt, or were classified as having some other improper use of the belt, and the small fraction of FARS CRS users who had grossly improper use, such as not having the CRS restrained with a seat belt still suggest little protective effect from use of CRS relative to seat belts in an unadjusted cohort analysis; however, because younger children are associated with higher risks of death and higher CRS use, an adjusted analysis suggests a 38% reduction in risk (RR = 0.62, 95% CI = 0.42, 0.93). The method of Levitt and Porter still suggests no protective effect for CRS, in both the unadjusted and adjusted analyses, similar to the results obtained from a standard cohort analysis using FARS data alone. The point estimate in matched case-control analysis suggests a protective effect for CRS-restrained children, although the limited sample size (288 vehicles containing 503 children) provides limited power to detect modest differences, as evidenced by the relatively wide confidence intervals. The robust relative risk estimator suggests a 22% reduction in risk of death for CRS restrained children versus belt-restrained children, although this difference does not quite reach statistical significance (RR = 0.78 95% CI = 0.63, 1.03).

Table 6 shows the estimated robust relative risk estimator, stratified by propensity score quintile. This table suggests a tendency for all children aged 2 through 6 to receive benefit from the use of CRS instead of a seat belt regardless of how likely they are to actually be restrained in one, although all propensity strata are limited in their ability to detect a significant difference because of small sample sizes.

5. DISCUSSION

The effectiveness (or lack thereof) of CRS relative to seat belts for children aged 2–6 would ideally be ascertained using an unobservable population: those children involved in a crash in which they would have died had they been restrained in a CRS, a belt, or in either type – that is, the total population of passenger vehicle crashes in which a restrained child aged 2–6 was in the vehicle from which the crashes survivable under either restraint type have been removed. If restraint use were randomized in the population, a standard case-control or cohort analysis would consistently estimate CRS effectiveness in this population. However, restraint use is likely not effectively randomized in the population with respect to driving behavior, so such an analysis can overstate or understate the effectiveness of a restraint. Using data from fatal crashes only might appear to solve this dilemma, but it will underestimate the effectiveness of CRSs if: a) CRSs are indeed effective relative to seat belts and b) if there is a positive correlation between CRS use and good driving behavior. Our analysis suggests that both types of biases may be present when obtaining estimates of CRS effectiveness relative to seat belts from either the FARS census of crashes with one or more fatalities or a FASS-NASS combination of datasets that are representative of all towaway crashes. We take advantage of the fact that a census of fatal crashes is obtained to develop a simple propensity score method to counter the selection bias approaches inherent in both analyses. This results in estimates of restraint effectiveness that may more accurately reflect the reductions in mortality risk that accrue from the use of the CRS itself, rather than from

the type of driver who chooses to restrain the children in their car in a CRS, and thus the reductions in mortality risk that will accrue as CRS use spreads throughout the remainder of the driving population. Our conservative robust risk ratio estimator suggests that the risk of mortality for children aged 2 through 6 restrained in a CRS relative to those restrained in seat belt is in the range of an increase of 10% to a decrease of 32%. This is less effective than the increase of 2% to a decrease of 54% estimated by a standard cohort analysis, but more effective than the estimate of an increase of 37% to a decrease of 1% that is obtained using FARS data alone. It appears that failing to observe the "children that did not die" can underestimate CRS effectiveness; younger children are overrepresented in the FARS data, and they are more likely to be restrained in a car seat, or, put in terms of the propensity score, relying on the FARS data alone to estimate (2) causes an underestimation of high propensity score and an overestimation of low propensity scores. On the other hand, using a standard cohort analysis can yield overestimates of restraint effectiveness because of the failure to account for potentially higher rates of survivable crashes among CRS users.

The Levitt and Porter approach appear to reasonably estimate the distribution of potential outcomes in vehicles where someone else in the crash dies – and thus restraint effectiveness in the top half of Table 1, but because it does not consider the distribution of potential outcomes in otherwise non-fatal (i.e., no one in the "other" vehicle died) crashes – restraint effectiveness in the bottom half of Table 1, it appears to suffer from the same bias toward the null as a standard FARS-only analysis. The matched-pairs analysis appears to have overcome this limitation by matching on vehicle; its estimate of a protective effect for CRSs that are not severely misused is similar to that of our robust causal estimator.

Our propensity score approach is not without limitations. Results can be sensitive to the choice of the propensity score model; in this analysis, a propensity score estimated using additional main effects for seat row, model year, and driver age had a substantial impact on the relative risk estimate (RR = 0.89). As Rosenbaum (1998) notes, even if the true propensity score is known, only "overt" biases can be eliminated. Thus the propensity score is not a perfect substitute for true randomization, which will asymptotically balance both observed and unobserved confounders.

Another limitation of any analysis focusing on fatality outcomes is that it ignores the large amount of morbidity that is likely prevented by the use of CRSs instead of seat belts in the 2 through 6 year-old population. "Seat belt syndrome" has been well understood for five decades as a special risk for restrained young children in passenger vehicle crashes (Agran et al. 1987; Garrett and Braunstein 1962; Kulowski and Rost 1956). Arbogast et al. (2004) showed that children aged 1 through 3 in forward-facing car seats reduced their risk of injury by 71% over children in seat belts. Durbin et al. (2003) found that children aged 4 through 7 in belt-positioning booster seats reduced their risk of injury by 59% over children in seat belts. Also, exposure to CRS and belt in the FARS database is measured by police report, which is subject to measurement error due to potential bias in police reporting of restraint. If this error is essentially random, the resulting estimates of CRS effectiveness will be biased toward the null, suggesting that the CRS protective effect will be underestimated. Non-differential bias – e.g., police being more likely to report a fatally-injured child as restrained

in a CRS when they were in a belt than a nonfatal crash victim, or vice-versa – may lead to either over-or underestimation of CRS effectiveness.

Some of the methods given here did not find even marginally protective effects of CRS relative to seat belts. However, this lack of protective benefit contradicts the known biomechanical properties of child restraint systems. Optimal performance of restraint systems depends upon an adequate fit between the restraint system and the occupant at the time of the crash. Child restraint systems are designed to reduce risk of ejection during a crash, better distribute the load of the crash through structurally stronger bones rather than soft tissues, limit the crash forces experienced by the vehicle occupant by prolonging the time of deceleration, and potentially limit the contact of the occupant with intruding vehicle structures. The analyses in this manuscript using a full cohort of fatal and non-fatal crashes generally suggest reductions in risk of death on the order of 20%. Given the relatively small number of fatal outcomes in this age range during the time period of study, the question of whether CRSs are truly effective relative to seat belts above and beyond associations with the driving behaviors of drivers who use them remains open to some degree. Future efforts to assess restraint effectiveness might consider instrumental variables approaches (Bowden and Turkington 1984) or other causal modeling techniques such as principle stratification (Frangakis and Rubin 2002) to reduce selection bias that may be inherent in observational crash data.

Acknowledgments

References

Agran P, Dunkle D, Winn D. Injuries to a Sample of Seatbelted Children Evaluated and Treated in a Hospital Emergency Room. Journal of Trauma. 1987; 27:58–64. [PubMed: 3806714]

Arbogast KB, Durbin DR, Cornejo RA, Kallan MJ, Winston FK. An Evaluation of the Effectiveness of Forward Facing Child Restraint Systems. Accident Analysis and Prevention. 2004; 36:585–589. [PubMed: 15094411]

Bowden, RJ., Turkington, DA. Instrumental Variables. Cambridge, UK: Cambridge University Press; 1984. MR0798790

Dubner SJ, Levitt SD. Freakonomics; The Seat-Belt Solution. New York Times Magazine. Jul 10.2005 :20. 2005.

Durbin DR, Elliott MR, Winston FK. Belt-positioning Booster Seats and Reduction in Risk of Injury Among Children in Vehicle Crashes. Journal of the American Medical Association. 2003; 289:2835–2840. [PubMed: 12783914]

Elliott MR, Kallan MJ, Durbin DR, Winston FK. Effectiveness of child safety seats vs seat belts in reducing risk for death in children in passenger vehicle crashes. Archives of Pediatric and Adolescent Medicine. 2006; 160:617–621.

Evans L. Double Pair Comparison – a New Method to Determine How Occupant Characteristics Affect Fatality Risk in Traffic Crashes. Accident Analysis and Prevention. 1986; 18:217–227. MR0853966. [PubMed: 3730096]

Frangakis CE, Rubin DB. Principal Stratification in Causal Inference. Biometrics. 2002; 58:21–29. MR1891039. [PubMed: 11890317]

Garrett JW, Braunstein PW. The Seat Belt Syndrome. Journal of Trauma. 1962; 2:220–238. [PubMed: 13897066]

Holland PW. Causal Inference, Path Analysis, and Recursive Structural Equation Models. Sociological Methodology. 1988; 1988:449–484.

Joffe MM, Ten Have TT, Feldman HI, Kimmel SE. Model Selection, Confounder Control, and Marginal Structural Models: Review and New Applications. The American Statistician. 2004; 58:272–279. MR2109415.

Kulowski K, Rost W. Intra-abdominal Injury from Safety Belts in Auto Accidents. Archives of Surgery. 1956; 73:970–971. [PubMed: 13372023]

Leon AC, Mueller TI, Solomon DA, Keller MB. A Dynamic Adaptation of the Propensity Score Adjustment for Effectiveness Analyses of Ordinal Doses of Treatment. Statistics in Medicine. 2001; 20:1487–1498. [PubMed: 11343369]

Levitt, Steven D. Evidence that Seat Belts are as Effective as Child Safety Seats in Preventing Death for Children Aged Two and Up. 2005. NBER Working Paper No. W11591Available at SSRN: http://ssrn.com/abstract=800446

Levitt S, Porter J. Sample Selection in the Estimation of Air Bag and Seat Belt Effectiveness. The Review of Economics and Statistics. 2001; 83:603–615.

National Highway Traffic Safety Administration. National Automotive Sampling System (NASS) Crashworthiness Data System. US Department of Transportation; Washington, DC: 1997. http://www-nass.nhtsa.dot.gov/NASS/cds/AnalyticalManuals/aman1997.pdf

National Highway Traffic Safety Administration. National Automotive Sampling System (NASS) General Estimates System, US Department of Transportation; Washington, DC: 2000. http://www.nhtsa.dot.gov/people/ncsa/pdf/GESmanual88-99.pdf

National Highway Traffic Safety Administration. FARS Analytic Reference Guide, 1975-2002. US Department of Transportation; Washington, DC: 2005. http://ftp.nhtsa.dot.gov/FARS/FARS-DOC/USERGUIDE-2002.pdf

Robins JM. Association, Causation, and Marginal Structural Models. Synthese. 1999; 121:151–179. MR1766776.

Robins JM, Merman MA, Brumback B. Marginal Structural Models and Causal Inference in Epidemiology. Epidemiology. 2000; 11:550–560. [PubMed: 10955408]

Rosenbaum, PR. Propensity score. In: Armitage, P., Colton, T., editors. Encyclopedia of Bio-statistics. Chichester, UK: Wiley; 1998.

Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70:41–55. MR0742974.

Rubin DB. Estimating Causal Effects in Randomized and Non-randomized Studies. Journal of Educational Psychology. 1974; 66:688–701.

Rubin, Donald B. Bayesian inference for causal effects: The role of randomization. The Annals of Statistics. 1978; 6:34–58. MR0472152.

Rubin DB. Comment on Neyman 1923 and Causal Inference in Experiments and Observational Studies. Statistical Science. 1990; 52:472–480. MR1092987.

Rubin DB. Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies. Journal of the American Statistical Association. 1979; 74:318–328.

Rubin DB, Thomas N. Matching Using Estimated Propensity Scores: Relating Theory to Practice. Biometrics. 1996; 52:249–264. [PubMed: 8934595]
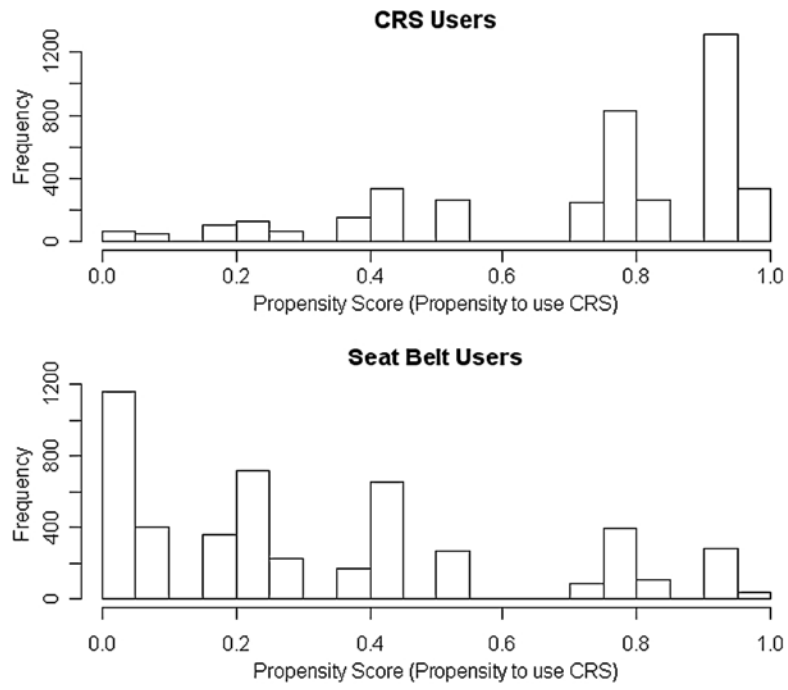
Figure 1.
Propensity score by restraint type.

Table 1

Illustration of a population of child outcomes in passenger vehicle crashes if potential fatality status $Y(T)$ for a restrained child given that s/he is restrained in type T could be observed. Shaded cells are not observable in the FARS dataset. For illustrative purposes, the simplifying assumption is made is that child restraint systems are never harmful ($Y(1) \leq Y(0)$). (RR = relative risk). "True RR" is relative risk estimated as the risk of death if all children were restrained in CRS relative to the risk of death if all children were in belts among the subset of "potentially fatal" crashes. "Cohort RR" ignores conditioning on survivability of crash and computes relative risk using observed data in sample (either FARS only or from a representative sample of the whole population)

Crash Type	Child dies if in CRS	Child dies if in seat belt	Others die in crash	(A): No association between restraint and fatal crash		(B): Children in CRS ~1/5th as likely to be in crash where others die		(C): Children in CRS ~5 times as likely to be in crash where others die	
				Child in CRS	Child in belt	Child in CRS	Child in belt	Child in CRS	Child in belt
Always Fatal	Yes $Y(1)=1$	Yes $Y(0)=1$	Yes $W=1$	1000 (a)	1000 (g)	1000 (a)	1000 (g)	1000 (a)	1000 (g)
Fatal in belt only	No $Y(1)=0$	Yes $Y(0)=1$	Yes $W=1$	100 (b)	100 (h)	100 (b)	100 (h)	100 (b)	100 (h)
Always Survive	No $Y(1)=0$	No $Y(0)=0$	Yes $W=1$	10000 (c)	10000 (i)	10000 (c)	10000 (i)	10000 (c)	10000 (i)
Always Fatal	Yes $Y(1)=1$	Yes $Y(0)=1$	No $W=0$	10 (d)	10 (j)	50 (d)=5(j)	10 (j)	2 (d)=(j)/5	10 (j)
Fatal in belt only	No $Y(1)=0$	Yes $Y(0)=1$	No $W=0$	100 (e)	100 (k)	500 (e)=5(k)	100 (k)	20 (e)=(k)/5	100 (k)
Always Survive	No $Y(1)=0$	No $Y(0)=0$	No $W=0$	100000 (f)	100000 (l)	500000 (f)=5(l)	100000 (l)	20000 (f)=(l)/5	100000 (l)

True RR

$$\frac{[a+d+g+j]/[a+b+d+e+g+h+j+k]}{[a+b+d+e+g+h+j+k]/[a+b+d+e+g+h+j+k]}$$

A: $\dfrac{2020/2420}{2420/2420} = .83$ B: $\dfrac{2060/2860}{2860/2860} = .72$ C: $\dfrac{2012/2332}{2332/2332} = .86$

Cohort RR: FARS only

$$\frac{[a+d]/[a+b+c+d]}{[g+h+j+k]/[g+h+i+j+k]}$$

A: $\dfrac{1010/1110}{1210/1120} = .84$ B: $\dfrac{1050/11150}{1210/1120} = .87$ C: $\dfrac{1002/11102}{1210/11210} = .84$

Cohort RR: Population cohort

$$\frac{[a+d]/[a+b+c+d+e+f]}{[g+h+j+k]/[g+h+i+j+k+l]}$$

A: $\dfrac{1010/111210}{1210/111210} = .83$ B: $\dfrac{1050/511650}{1210/111210} = .19$ C: $\dfrac{1002/31122}{1210/111210} = 2.95$

Table 2

Sample sizes used for analyses, by source of data

	FARS		NASS-CDS	
	N of Vehicles	**N of Children**	**N of Vehicles**	**N of Children**
Observational Cohort/Robust Causal	6391	7816	982	1436
Levitt and Porter Restricted Sample	3204	3940	—	—
Matched Case-Control	228	503	—	—

Table 3

Child occupant and crash characteristics, 1998–2003. Data are presented as weighted for NASS-CDS % (unweighted n in parentheses)

Characteristic	FARS	NASS-CDS	All
Restraint Use			
Seat Belt	53.7[*] (4203)	54.2 (694)	54.2 (4307)
Child Restraint System	46.2[*] (3613)	45.8 (742)	45.8 (4945)
Age			
2 Years	24.4 1911	21.1 (334)	21.2 (2245)
3 Years	21.8 (1705)	17.4 (279)	17.4 1984
4 Years	20.3 (1585)	19.1 (290)	19.1 1875
5 Years	16.3 (1344)	15.0 (271)	15.0 (1615)
6 Years	15.4 (1271)	27.4 (262)	27.3 (1533)
Seating Position			
Front	17.3 (1351)	15.7 (200)	15.7 (1571)
Rear	82.7 (6465)	84.3 (1236)	84.3 (7701)
Vehicle Type			
Passenger Car	53.2 (4161)	67.8 (846)	67.7 (5007)
Pickup Truck	11.0 (859)	5.6 (83)	5.6 (942)
Van/Minivan	17.8 (1390)	17.2 (299)	17.2 (1689)
Sports Utility Vehicle	18.0 (1406)	9.4 (208)	9.5 (1614)
Model Year			
<1990	19.1 (1492)	15.5 (258)	15.5 (1750)
1990–1993	21.4 (1674)	23.4 (280)	23.4 1954
1994–1997	29.1 (2274)	33.8 (420)	33.8 (2694)
1998–2004	30.4 (2376)	27.3 (478)	27.3 2854
Driver Age			

Characteristic	FARS	NASS-CDS	All
<20	8.2 (560)	4.0 (43)	4.0 (603)
20+	91.8 (7256)	96.0 (1393)	96.0 (8649)

*Includes 92 subjects with severe belt misuse and 162 subjects with severe CRS misuse.

Table 4

Child restraint system use (vs. seatbelt) by occupant and crash characteristics, within weighted propensity score quintiles. P-value under null hypothesis of no association between CRS use and the characteristic within each stratum. Data are presented as weighted for NASS-CDS % (unweighted n in parentheses)

Characteristic		PS 1	p	PS 2	p	PS 3	p	PS 4	p	PS 5	p
Child Age	2	(0)		(0)		(0)		91.0 (591)		92.8 (1654)	
	3	(0)		(0)		47.8 (195)		79.0 (1789)		(0)	
	4	7.2 (189)	.35	33.0 (324)	.53	45.2 (1362)	.89		.11		—
	5	0.5 (169)		24.8 (1446)		(0)		(0)		(0)	
	6	4.2 (1533)		(0)		(0)		(0)		(0)	
Driver Age	<20	0.1 (117)		71.2 (123)		40.4 (84)		96.5 (144)		75.3 (135)	
	20+	4.5 (1774)	.22	23.6 (1547)	.15	45.4 (1473)	.87	80.7 (2236)	.30	93.9 (1519)	.23
Seated in Front Row	Yes	3.0 (575)		0.7 (264)		39.0 (298)		88.0 (303)		79.0 (111)	
	No	4.4 (1316)	.62	29.2 (1506)	.03	47.4 (1259)	.55	78.9 2077	.50	93.6 (1543)	.27
Vehicle Type	Passenger Car	2.9 (781)		27.9 (861)		42.0 1019		77.2 1070		93.5 (1276)	
	Pickup Truck	5.6 (544)	.51	(0)	.88	47.8 (195)	.51	86.6 (203)	.65	(0)	.60
	Van/Minivan	8.1 (291)		24.2 (298)		64.7 (343)		82.1 (379)		88.5 (378)	
	Sports Utility Vehicle	0.7 (275)		21.9 (611)		(0)		89.0 (728)		(0)	
Model Year	<1990	3.2 (341)	.75	9.6 (315)	.21	42.5 (330)	.30	48.0 (417)	.27	92.2 (347)	.73

Characteristic		PS 1	PS 2	PS 3	PS 4	PS 5
	1990–1993	3.1 (369)	17.0 (361)	78.9 (371)	73.3 (452)	89.0 (401)
	1994–1997	2.4 (547)	41.0 (507)	31.1 (456)	87.7 (710)	94.0 (474)
	1998–2004	7.0 (634)	33.0 (587)	41.3 (400)	88.3 (801)	95.3 (432)
Driver Gender	Male	5.2 1084	24.4 (853)	38.9 (767)	76.4 (1247)	95.6 (832)
	Female	3.9 (807)	26.3 (917)	47.4 (790)	82.9 (1133)	91.2 (698)
		.65	.81	.57	.54	.43
Overall		4.2 1891	25.8 (1770)	45.4 (1557)	80.8 (2380)	92.8 (1654)

Table 5

Relative risk measures, using various observational and causal methods discussed in the text. Analyses restricted to towaway crashes only (95% confidence intervals in parentheses)

Population Cohort	Observational Cohort		Levitt and Porter Restricted Sample Method		Matched Case-Control	Robust Causal (RR_2)
	Unadjusted	Adjusted	Unadjusted	Adjusted		
Full population (FARS and NASS)	1.02 (.72,1.45)	0.69 (0.46,1.02)	N/A	N/A	N/A	0.83 (0.68.1.10)
Full population excluding severe misuse	0.92 (0.65,1.31)	0.62 (0.42,0.93)	N/A	N/A	N/A	0.78 (0.63.1.03)
FARS only	1.01 (0.88,1.15)	1.16 (0.99,1.37)	1.02 (0.82,1.28)	1.14 (0.85,1.51)	0.93 (0.61,1.43)	N/A
FARS only excluding severe misuse	0.92 (0.80,1.05)	1.05 (0.89,1.26)	0.94 (.75,1.20)	1.01 (0.75,1.36)	0.79 (0.49,1.25)	N/A

Table 6

Causal relative risk measures using full population: overall, by propensity score quintile (PSQ), and by age

Population Cohort	Robust Causal (RR_2)
PSQ 1	0.82 (.25,1.98)
PSQ 2	0.69 (.47,0.96)
PSQ 3	0.99 (0.73,1.37)
PSQ 4	0.89 (0.68,1.16)
PSQ 5	0.78 (0.45,1.27)