METHODOLOGICAL SHORTCOMINGS AND OPPORTUNITIES IN AUTOMOBILE SAFETY RESEARCH AS WE ENTER THE AUTONOMOUS ERA

Vikram Maheshri
University of Houston

Clifford Winston*
Brookings Institution

Abstract. We argue that the established and varied research on highway safety suffers from fundamental methodological shortcomings. Researchers face the dilemma of either identifying causal influences on automobile accidents or asking limited policy-oriented questions about auto safety. We develop a theoretical framework to clarify the limitations of the three main approaches taken in automobile safety research: the use of controlled environments, disaggregated data, and aggregated data. We illustrate these limitations in the context of the vast empirical literature that has sought to assess the effectiveness of seatbelt use in reducing fatal accidents. We briefly discuss promising advances in computation and data that may help improve the credibility of automobile safety research as we enter an era of vehicle autonomy.

May 2026

***We are grateful to Fred Mannering, Jia Yan, and a referee for helpful comments.**

1. Introduction

US policymakers and automakers have long prioritized improving automobile safety to reduce accidents and fatalities. Policymakers have spent hundreds of billions of dollars to improve the safety of roadways, expand and modernize traffic enforcement, and conduct public safety campaigns to discourage speeding and driving under the influence of alcohol and drugs. Automakers have strengthened vehicles' structures to provide greater resistance to crashes, improved braking and steering, installed occupant safety devices, and recently begun to make autonomous vehicle safety features available. As a result of safety improvements in the driving environment and automobiles, the US experienced a sustained 3% annual decline, on average, in the rate of automobile fatalities from 1920 to 2010. But since 2010, the US has not reduced the rate of automobile fatalities while experiencing a slight increase since COVID began in 2020 (see figure 1).[1]

Although it is unreasonable to expect that the US would maintain the same rate of safety improvement indefinitely, the recent lack of improvement is cause for concern because the annual cost of automobile accidents, including fatalities, the loss of quality of life, vehicle damage, and disruption of road travel, currently amounts to more than $1 trillion annually (TRIP, 2023). Such a large social cost should encourage the research community, including economists and transportation engineers, to improve our understanding of the causes of highway accidents and to provide evidence of the potential for new public policies to reduce this cost.

However, we argue in this paper that the prevailing micro-based research methodologies in highway safety—specifically those relying on controlled environmental studies and disaggregated police data—suffer from inherent and, in practice, fatal methodological flaws that preclude the identification of the causal parameters that would actually be of use to researchers trying to understand the causes of auto safety in real-world contexts or policy-makers trying to foster a safer driving environment. While research using aggregated data avoids these specific selectivity biases, it is necessarily constrained in the scope of questions it can address.

In controlled environments, the absence of physical and financial risk nullifies the fundamental behavioral trade-offs—such as the Peltzman (1975) offsetting effects—that define real-world driving. In disaggregated police data, the literature is plagued by a foundational

---

[1] Estimates of the rate of change of automobile fatalities in the US from 1920 to 2010 and from 2010 to 2023 are based on annual fatalities data from the US Department of Transportation.

selectivity bias: by conditioning on a crash, researchers are blinded to the non-events that represent the true successes of safety technology and policy.

We recognize this is a stringent judgment of the field, yet we believe it is justified by three observations. First, we find no empirical work in these domains that successfully reconciles the endogeneity of driver risk with the technical performance of the vehicle. Second, while it is sometimes argued that safety research serves purposes beyond causal policy analysis—or that crash-conditioned work is nonetheless useful—this defense remains an empty vessel. We cannot cite a single paper that acknowledges its findings are non-causal while providing a persuasive framework for how such findings can reliably guide multi-billion dollar safety interventions. Without such a framework, non-causal associations offer a false sense of certainty regarding safety trends that may, in practice, justify needlessly costly, ineffective, or even counterproductive policies. Finally, the persistence of these approaches suggests a research program that has yet to engage in a necessary reckoning with its analytical limitations and its reliance on biased data. We hope this paper serves as that catalyst.

The analytical challenge in studying automobile safety arises from the fact that driving is a highly complex and dynamic behavior. Research approaches that ignore the behavioral aspects of driving will likely confuse the true causal influences of auto safety with the behavioral responses to vehicles and the driving environment that mediate those influences. Accordingly, the effects estimated using those approaches will lack external validity to the actual environments that drivers inhabit and policy-makers shape.

Observed outcomes (accidents) are caused by multiple decisions made by drivers, some of which are made even before they begin their automobile trip, including where, when, under what road conditions, in which vehicle, and in what physical and mental state they choose to drive. Other influential decisions on accidents are made by drivers during the trip, such as the speed and aggressiveness while driving. Those decisions in turn have countless determinants that are potentially highly correlated to one another and mostly unobservable, with unobserved risk preferences being the most important influence because they sort drivers into riskier or safer driving environments and behaviors.

To assess the impact of a potential influence on auto safety, we must therefore observe how it operates when accidents do and do not occur. For example, if we wish to estimate the impact of seat-belts on auto-safety, we must observe how the health of drivers who wear seat-belts is affected in both situations. Without observing drivers who wear seat-belts and who do not get into accidents, we have no counterfactual from which we can draw causal inferences. Given the behavioral influences on automobile safety, the fact that accidents do not occur is generally systematic; thus, the bulk of empirical auto-safety research, which relies on data that are generated only when an accident occurs and summarized in police accident reports, is particularly susceptible to the lack of an appropriate counterfactual. Alternative approaches intended to circumvent identification problems are similarly susceptible to systemic biases.[2]

Drivers also may adjust their behavior in response to a potential influence on auto safety. For instance, a driver whose vehicle is equipped with airbags may drive more aggressively than a driver whose vehicle is not equipped with airbags. Hence, even a comparison of the condition of drivers in accidents in vehicles with and without airbags will not constitute a valid counterfactual for causal inference.

These problems in automobile safety research have not been apparent because researchers have not prioritized formulating a theory of motorists' behavior that acknowledges the various complex decisions and influences that lead to accidents and that could guide the specification and identification of an econometric model of driving. By not taking this systematic approach, researchers have been prone to blindly diving into different data sets generated from controlled testing environments, police reports on individual motorists' accidents, and various administrative and private sources that aggregate accidents. But the use of each type of data faces distinct

---

[2] For example, Evans (1986) introduced a double-pair comparison to control for crash severity by comparing the fatality risk of a belted driver to an unbelted passenger in the same vehicle. However, by design this approach only can identify effects conditional on the occurrence of a crash and hence fails to account for endogenous behavioral offsets—specifically, the high probability that a driver's risk-taking behavior is a function of their perceived safety while belted. (The inverse issue would arise when comparing fatality risks of unbelted drivers and belted passengers.) Nevertheless, many studies have adopted this approach. The induced exposure method (Thorpe, 1964; Haight, 1970) bifurcates two-vehicle crashes into at-fault and not-at-fault (innocent) categories, assuming that the latter group constitutes a random, representative sample of the driving population at a given time and place. For example, if teenagers represent 10% of innocent victims but 40% of at-fault drivers, the method identifies teenagers as a high-risk group. However, innocence in a police report is not an exogenous draw; it is subject to a significant selection bias in exposure. Highly skilled and attentive drivers are more likely to avoid collisions through superior defensive maneuvers. Consequently, the innocent driver pool is systematically skewed toward lower-skill or more distracted individuals who fail to avoid the crash, rendering the random sample assumption invalid. Finally, Naturalistic Driving Studies employ in-vehicle cameras to observe pre-crash dynamics in real-time. These studies are compromised by the Hawthorne Effect, as drivers may alter their behavior under observation, and a profound selection bias from endogenous participation, as individuals willing to be continuously monitored are unlikely to be representative of the high-risk or marginalized driving populations most relevant to safety policy.

limitations that depend on the specifics of the research question, features of the data, and institutional details of the research settings. The controlled environment and disaggregate approaches have generally not been able to obtain credible causal explanations of the determinants of highway accidents and accurate estimates of the potential effects of government safety policies to reduce accidents. Although the aggregate approach can obtain credible causal explanations of the effects of certain policies, it is limited in its scope because it requires quasi-experimental variation that is not always available.

In what follows, we outline a dynamic theoretical framework of driver behavior that encompasses essentially all of the important automobile safety questions that are relevant to policymakers and researchers. This permits a consistent and rigorous comparison of the various empirical approaches that have been used by researchers and offers much needed clarification of the assumptions that underly each of them. In many cases, the assumptions that researchers are forced to make to achieve identification of their models are implausible, or the empirical findings from the approaches do not coincide with the effects that researchers and policymakers are actually interested in estimating. As an illustrative exercise, we apply the framework to clarify the weaknesses of current safety research methodologies in the context of answering an important, long-studied, and policy-relevant issue: How effective are seatbelts in reducing fatal accidents?

Taking an optimistic look to the future, we stress that autonomous vehicles have the potential to make enormous improvements in safety that could effectively eliminate vehicle accidents by neutralizing the risks posed by the nation's most dangerous drivers (Winston, Yan, and Associates, 2024). Indeed, credible evidence exists that even low levels of autonomous safety technology can reduce accidents and fatalities (Maheshri, Winston, and Yu, 2025). In addition, as fully autonomous vehicle technologies are adopted, the potential for identifying credible causal safety effects using disaggregated data becomes significantly more viable. The primary hurdle in current research—the contamination of data by unobserved driver risk preferences—is mitigated because the autonomous system, rather than the human driver, becomes the primary tactical agent as the driver is transformed from an endogenous source of behavioral variation into an exogenous passenger.[3] Moreover, the wealth of data produced by autonomous driving systems has the

---

[3] Of course, motorists with autonomous vehicles will still make the (non-random) choices of where and when to make their trips, which could raise identification issues when assessing the efficacy of autonomous driving technologies. However, given that these driving technologies are likely to be highly effective when they are widely adopted, these identification issues may be of minor importance.

potential to offer an unprecedented window into the travel behavior of *all* drivers, not simply those who suffer accidents.

We conclude by drawing two important lessons for the research community and policymakers. First, because the disaggregated approach using data from police accident reports has become the dominant approach for analyzing automobile safety, the broader research community must come to terms with the fact that it is unable to obtain well-identified causal estimates of the determinants of automobile accidents that are useful for safety policy. We briefly discuss the progress that does seem possible thanks to computational and data advances.

Second, as the United States adjusts to a new environment of autonomous vehicle technologies, policymakers should be aware of the tendency for flawed approaches to produce inflated estimates of their effectiveness when developing policy. Importantly, they should avoid past mistakes by not rushing to seize a misguided opportunity to improve highway safety by prematurely requiring automakers to install autonomous safety features in their new vehicles on the basis of flawed or speculative estimates of the direct and external benefits of those features.

2. A Theoretical Framework to Study Driving Behavior

Identification of econometric models arises from a combination of data and assumptions regarding the data generation process that is sufficiently informative to allow for inference. Parametric models of automobile safety contain parameters that correspond to the causal effects of a particular variable or variables. The thrust of our critique is that much of the published empirical findings in automobile safety research is based on econometric models that are not identified. The lack of identification in safety models does not originate from the samples per se that are used to analyze the determinants of accidents. Rather, the problem originates from a lack of a plausible theory of driver behavior. Without such a theory, the assumptions that allow for a model's parameters to be interpreted as causal effects are implausible and potentially incompatible with the data generation process itself. Because we are not aware of any empirical research in the safety literature that starts with a plausible theoretical framework to understand driving behavior and to determine the conditions under which an empirical model would be identified, we begin our formal critique of this literature by developing such a framework.

Driving is a complex dynamic activity because many agents (drivers) continuously update their positions on a road, which changes their exposure to other agents on the road and to roadway conditions. Previous work, for example, Tscharaktschiew (2020) and Yang et al. (2015), has modeled drivers' speed choices in a non-cooperative setting to obtain a traffic network equilibrium. We consider the simple case in which a fixed set of n drivers travel along the same highway route and choose their travel speed. The speed affects not only drivers' (future) position along the road, but it also affects the likelihood of a potentially significant cost; that is, the probability of and accident severity from being involved in an automobile accident.

The simple case that we analyze here raises a host of empirical issues that are not addressed in the empirical literature, such as treating route choice as an endogenous dynamic decision and endogenizing pre-trip decisions such as the choice of vehicle or time of day to drive. Adding further real-world complexity to our framework, such as motorists' choices of the extent and level of insurance coverage and how those choices may affect driving behavior, only exacerbates the problems with current empirical approaches to automobile safety.

The existence of multiple choices implies that drivers must make tradeoffs involving safety, such as taking a faster but more hazardous route, a more convenient but risky time to drive, and so on. Because drivers generally make these choices to maximize their self-interest subject to various uncertainties and exogenous variables, it is appropriate to formulate driving behavior as a dynamic optimization problem where the objective is expected utility and the state variable is the position of all drivers on the road.[4]

Formally, we denote the position of each driver i at time t as p_{it}, which we coalesce into an $n \times 1$ vector P_t. (Hereafter, upper-case variables correspond to the n vector of their lower case scalar counterparts.) The choice variable of each driver i is the travel speed at time t, which we denote as s_{it}. We characterize the driver's expected utility optimization problem with the following Bellman equation:

---

[4] While we formulate driving behavior within a dynamic expected utility (EU) framework, we recognize that alternative behavioral models, such as prospect theory (Kahneman and Tversky, 1979), offer nuanced insights into risk perception, such as the overweighting of small-probability events. However, the identification failure that we highlight does not stem from the *functional form* of the driver's utility, but from the endogeneity of risk. Whether a driver is a perfectly rational agent or one subject to behavioral bias, a fundamental feedback loop remains: agents adjusts their behavior in response to the perceived safety of the environment. Indeed, introducing behavioral complexity might exacerbate, rather than resolve, the identification problem by adding further unobserved latent variables to an already unidentified system.

$$V_i(P_t) = \max_{s_{it}} \left\{ \begin{array}{l} \underbrace{u_i(p_{it} + s_{it}) - \sum_{j=1}^{J} E_{it}\left[\pi_{it}^{j}(P_t, S_t, X_t)\right] \times c_{ij}}_{\text{flow utility}} + \\ \underbrace{\left(1 - \sum_{j=1}^{J} E_{it}\left[\pi_{it}^{j}(P_t, S_t, X_t)\right]\right)}_{\text{Prob. of not getting in accident}} \cdot \underbrace{V_i(E_{it}[P_{t+1}(S_t)])}_{\text{Continuation Value}} \end{array} \right\}, \tag{1}$$

where u_i is driver i's utility function, which increases with their position along the road, and π_{it}^{j} is the probability that driver i gets into an accident of severity j at time t, with severity ranging from vehicle damage only to a driver fatality. This probability is a function of the positions p and speeds s of all the drivers on the road as well as the relevant characteristics of the drivers and vehicles on the road and the roadway conditions at time t. We coalesce those variables in the matrix X_t. Finally, c_{ij} is the cost to driver i of getting in an accident of severity j.

The instantaneous flow utility in the equation that accrues to a driver is given by the utility from travelling an additional distance of s_{it} between periods t and $t + 1$ net of the expected costs of getting into an accident during that time. Conditional on not getting into an accident between periods t and $t + 1$, driver i obtains a continuation value given their beliefs of where all other vehicles will be on the road in period $t + 1$.

This stylized formulation of a driver's dynamic optimization problem captures three important and plausible features of driving that have critical implications for analyses of automobile safety: (1) Drivers form expectations of their safety based on the driving environment and their speed choices (π_{it}^{j} is a function of s_{it} and X_t); (2) Drivers form expectations of where other drivers will be, implicitly taking into account that those drivers face their own optimization decisions (π_{it}^{j} is a function of P_t and S_t); and (3) Drivers understand the decisions they make at any point in time may influence the decisions of other drivers at future times (P_{t+1} is a function of S_t). Because drivers' expectations may affect their likelihood of getting into an accident, those expectations must be accounted for by researchers if they wish to explain the determinants of accidents empirically.

The complexity of modelling a driver's problem grows because each driver faces her own analogous optimization problem in each period t. Hence, driving can be thought of as a dynamic game of incomplete information. The perfect Bayesian equilibrium of this game consists of a series of strategies, or mappings from the state space to the action space, which we denote as $S_{it}^{\star}(P_t)$, where * denotes an equilibrium value, accompanied by a specification of driver beliefs that satisfy

Bayes' Law. We therefore obtain the corresponding equilibrium accident probability functions $\pi_{it}^{j*}(P_t, X_t) = \pi_{it}^{j}(P_t, S_{it}^{*}(P_t), X_t)$.

3. Using the Framework to Analyze the Determinants of Automobile Safety

The major empirical questions of interest to automobile safety researchers can effectively be distilled into questions regarding the determinants of the accident probability, π_{it}^{j*}, in our framework. To connect this probability to several standard questions in the literature, we explicitly define the arguments in the matrix $X_t = (D_t, Z_t, R_t)$, where D_t is a matrix of driver socioeconomic characteristics, such as age and gender, Z_t is a matrix of vehicle characteristics, such as weight and horsepower, and R_t is a matrix of roadway characteristics, such as pavement condition and curvature.

Thus, the literature that seeks to inform policymakers and automakers by explaining how vehicle attributes affect auto safety, estimates $\frac{\partial \pi_{it}^{j*}}{\partial z_i}$ for driver i and an attribute z_i of her vehicle. The literature that seeks to inform policymakers by identifying the characteristics of drivers that contribute to risk, estimates $\frac{\partial \pi_{it}^{j*}}{\partial d_i}$ for a driver characteristic d_i. Finally, the literature that seeks to inform highway engineers by identifying how roadway conditions affect accident risk, estimates $\frac{\partial \pi_{it}^{j*}}{\partial r_t}$ for a roadway condition r_t.

Following standard econometric practice of estimating values of the mean parameters, researchers would estimate an average of those effects on accident risk over multiple configurations of different drivers and vehicles on the road during different time periods, thereby obtaining a treatment effect that does not vary over time t and is not a function of drivers' positions on the road P_t. This choice of aggregation raises immediate concerns of how the estimates of interest can be identified because P_t is a driver's state variable in her dynamic optimization problem and it *does* affect accident risk. This relationship constitutes a potential source of endogeneity that must be addressed to obtain consistent estimates of important influences on safety. Note that the identification problem does not depend on whether we use a random or nonrandom sample per se for the empirical analysis. Importantly, there is no plausible assumption that follows from a credible theory of motorists' behavior about the relationship between the endogenous variables and the error term that addresses the problem. We discuss various identification problems in the context of the different empirical approaches that have been taken in the safety literature.

4. Using the Framework to Assess Empirical Approaches in the Safety Literature

Economists and transportation engineers have taken three different empirical approaches to estimate the determinants of an accident probability: (1) a controlled environment approach that generates empirical observations from simulated accidents; (2) a disaggregate approach based on accident data generated by individual drivers and included in police accident reports; and (3) an aggregate approach based on accident data generated by travelers and aggregated to a geographic level, such as a state. Transportation engineers have primarily taken the first two approaches and economists have primarily taken the third approach. We clarify the identification problems that cause the controlled environment and disaggregate approaches to produce biased estimates. It does not appear that tractable methods are currently in use to circumvent the bias in those approaches. It is possible to circumvent the bias by taking an aggregate approach that limits the types of questions about accident safety that researchers can address.

We illustrate the limitations of each approach in the context of one of the most studied questions in the automobile safety literature: To what extent does wearing seatbelts reduce driving fatalities?

The Controlled Environment Approach

The controlled environment approach refers to a research design where researchers subject specific vehicles to simulated driving conditions and observe specific aspects of their safety performance. Crash tests and closed course observations are well-known examples of controlled environment approaches. Although policymakers are partial to this approach, in all likelihood because it resembles *randomized* controlled trials, which are broadly recognized as the gold standard of causal research (Kahane (2015)), the approach is far from being randomized.

The Lack of Randomness and Selectivity Bias. To illustrate the lack of randomness, which is a sufficient (but not necessary condition) to preclude identification, consider that a series of carefully conducted crash tests showed that the use of seatbelts reduced the risk of a fatal accident by 45% (e.g., Lave and Weber (1970)). Assuming that seatbelts are the vehicle attribute z_i of interest and that their effectiveness is determined by drivers' behavior to wear them, this finding corresponds to $\frac{\partial \pi_{it}^{j}}{\partial z_i}$ evaluated at a particular value of vehicles' positions, drivers' and vehicles' characteristics, and roadway conditions; that is, (P, X) corresponding to the details of the tests.[5] But this calculation differs from the calculation of interest in actual driving environments, $\frac{\partial \pi_{it}^{j*}}{\partial z_i}$, which is determined as the equilibrium choices of driver i that are evaluated at the equilibrium levels of (P^*, X^*). So, for instance, if safer drivers were more likely to wear seatbelts, then the calculation of $\frac{\partial \pi_{it}^{j}}{\partial z_i}$ based on the crash test would overestimate $\frac{\partial \pi_{it}^{j*}}{\partial z_i}$.[6] Alternatively, if wearing seatbelts make drivers feel safer and thus more willing to drive in potentially hazardous road conditions, such as during a snowstorm, then $\frac{\partial \pi_{it}^{j}}{\partial z_i}$ would underestimate $\frac{\partial \pi_{it}^{j*}}{\partial z_i}$.[7]

Of course, the potential selection bias from individuals' choice of seatbelt use is generally well-known, and some researchers may feel that it merits only a qualification. But we have shown that the bias is likely to be quite serious by showing how it arises in a plausible dynamic model of drivers' behavior and by clarifying that drivers' safety outcomes are based on a series of decisions that they make prior to and while driving. Moreover, we will show that this bias is quantitatively meaningful, as estimates of the effects of seatbelts obtained from this approach are substantially larger than estimates obtained from other empirical approaches. As pointed out, the decision they make while driving is their choice of speed s_{it} at time *t*, which determines their position p_{it} on the road at time *t*. The decisions they make prior to driving include the type of vehicle to buy, which determines vehicle characteristics Z_t, the kinds of behaviors to engage in, which determines driver characteristics D_t, and the roads they will traverse and when they will travel, which determines roadway characteristics R_t. As noted, these decisions are not made randomly or out of habit; they involve tradeoffs in real time and drivers will maximize their expected utility by making their preferred tradeoffs based on observed and unobserved influences.

---

[5] The 45% risk reduction technically corresponds to $\frac{\partial \log \pi_{it}^{j}}{\partial \log z_i}$. This elasticity can be recovered from the marginal effect $\frac{\partial \pi_{it}^{j}}{\partial z_i}$.

[6] Descriptive correlational evidence indicates that drivers who use seatbelts are less likely to engage in risky behaviors like speeding or impaired driving, based on observational data and crash statistics. See the "National Occupant Protection Use Survey" published as "Seat Belt Use in 2019—Overall Results" (Report No. DOT HS 812 821).

[7] In this case, drivers' behavior would be consistent with Peltzman's (1975) and Wilde's (1982) risk compensation hypothesis. Winston, Maheshri, and Mannering (2006) found that motorists' increase in risky driving behavior appeared to offset the technological effectiveness of early generation airbags and antilock brakes.

Is Selectivity Bias Fatal? By incorporating these considerations into a model of driver behavior, it becomes clear that estimates of seatbelt efficacy based on crash tests would be applicable to highway driving only if seatbelt use were determined independently of all other decision variables in our framework. Such an assumption is implausible and cannot be salvaged by a mere qualification.

That said, it could still be argued that although a controlled environment study cannot recover the policy parameters of interest, it remains informative for engineering design, trauma care, and the improvement of biomechanical protection. We contend that this defense is an empty vessel for two reasons. First, these fields treat the crash event as a static, exogenous given, ignoring the dynamic behavioral responses that determine whether a crash occurs at all. Biomechanical data may measure a vehicle's technical potential for protection, but it cannot predict its real-world safety performance when human risk-taking is endogenous.

Second, in the absence of a behavioral link, improvements in occupant protection can be actively misleading for policy; a vehicle that is safer in a lab may induce enough additional risk-taking to increase total societal fatalities. Consequently, while those studies may assist in hardware calibration, they are fundamentally incapable of addressing the causal questions that define public safety. To rely on them for policy is to mistake the physics of the collision for the economics of the accident.

The Disaggregate Approach

Researchers have attempted to circumvent one shortcoming of the controlled environment approach by using disaggregated data obtained from accidents involving actual drivers; thus, the disaggregate approach refers to a research design where researchers use observational, driver-level data to estimate the effects of various highway and vehicle characteristics and safety policies on accidents.

Our theoretical framework characterizes the behavior of all drivers regardless of whether they are involved in an accident. In addition, we do not assume that drivers who are involved in accidents do not differ from drivers who are not involved in accidents in terms of observed and unobserved influences on accidents. The determinants of the decisions that drivers make on the road are also correlated to observed influences on accidents, including speed choice, vehicle characteristics, some driver characteristics, and roadway conditions for their trip. Important examples of determinants of accidents that are not proxied, measured, or observed by the researcher are unobserved characteristics of the driver, such as their temperament and judgment.[8]

Lack of Randomness and Selectivity Bias. The immediate weakness of the disaggregate approach is that because researchers obtain data from police accident reports, they are forced to make the implausible assumption that drivers who are involved in accidents do *not* differ from drivers who are not involved in accidents in order to attach external validity to their results. This assumption merits more than a qualification and is much stronger than researchers realize. That is, it is assumed that the decisions drivers make prior to and while driving and their unobserved characteristics do not have different effects on drivers who get into accidents and on drivers who do not get into accidents. However, drivers effectively self-select to be included in accident reports by being involved in an accident; otherwise, they are not included in those reports.

Similar to researchers who take a controlled environment approach, researchers who take a disaggregate approach do not obtain findings based on a random sample. Researchers rarely acknowledge the implications of this problem and implicitly attempt to deal with it by effectively comparing drivers who get into accidents of different severities (i.e., fatality, serious injury, minor injury, or property damage only). But to avoid biasing parameter estimates, researchers must estimate an effect that captures *both* the marginal effect of getting into an accident and the conditional effect of the severity of that accident. Even if researchers attempt to use sophisticated econometric methodologies to, for example, control for motorists' heterogenous behavior, estimating the determinants of the severity of an accident conditional on an accident occurring is simply unable to address the fundamental identification problems that we have stressed here.

Again, a nonrandom sample is a sufficient but not a necessary condition to prevent identification. Indeed, even if a researcher could use a random sample, she would have to make strong and implausible assumptions to identify the marginal effect of getting into an accident and the conditional effect of the severity of that accident given that both outcomes are likely to be influenced by the same set of unobserved and unobserved variables.

Researchers have taken different empirical approaches to estimate disaggregate models, but they all essentially estimate the probability of getting in an accident of severity $j > 1$

---

[8] It could be argued that researchers use the number of motorists' speeding tickets as a proxy for risk preferences. For example, see Vertlib et al. (2023). Of course, that proxy tends to understate risk preferences because most incidents of speeding are not ticketed.

conditional on getting in an accident of any severity (denoted as π_{it}^{*}), where j = 1, if no accident occurred. We can write this probability as:

$$\pi_{it}^{j**} = \frac{\pi_{it}^{j*}}{\pi_{it}^{*}} \tag{2}$$

It is clear that $\pi_{it}^{j**} \neq \pi_{it}^{j*}$. Importantly, this fact dramatically diminishes the value of the entire empirical exercise if its ultimate purpose is to inform policymakers how highway safety could be improved.

Consider, for example, the policy question of the causal effect of seatbelt use on automotive fatalities. Denote seatbelt use with the binary variable z_i (where 1 indicates the use of a seatbelt; 0 otherwise). The effect that is identified in a disaggregate analysis can then be written as:

$$\text{Disaggregate Effect} = \left(\pi_{it}^{j*} | z_i = 1, A_i = 1\right) - (\pi_{it}^{*} | z_i = 0, A_i = 1)\ , \tag{3}$$

where the binary variable A_i is equal to 1 if the vehicle got in an accident. Policymakers are interested in what we call the true causal effect (TCE) of seatbelt use on fatalities, which can be expressed as:

$$\begin{aligned} \text{TCE} &= (\pi_{it}^{*} | z_i = 1) - (\pi_{it}^{*} | z_i = 0) \\ TCE &= \left(\pi_{it}^{j*} | z_i = 1, A_i = 1\right) \times (\pi_{it}^{*} | z_i = 1) - \left(\pi_{it}^{j*} | z_i = 0, A_i = 1\right) \times (\pi_{it}^{*} | z_i = 0) \end{aligned} \tag{4}$$

Even if we could perfectly estimate the probability of getting in an accident (π_{it}^{*}), we would be unable to use the estimates of the disaggregate causal effect in equation (3) to obtain the TCE unless we made the strong additional assumption that $(\pi_{it}^{*} | z_i = 1) = (\pi_{it}^{*} | z_i = 0)$. This assumption is highly implausible because it states that the likelihood that a person who wears a seatbelt gets in an accident is the same as the likelihood that a person who does not wear a seatbelt gets in an accident, which ignores that a person's propensity to wear a seatbelt is correlated to their attitude toward risk and, in turn, to their driving behavior. The assumption is further weakened because some drivers may adjust their behavior if they are wearing a seatbelt.

Although we have shown how the identification problem prevents one from determining the true causal effect on safety when the treatment variable is discrete, the same identification problem extends to the case when the treatment variable is continuous. For example, the problem arises in Anderson and Auffhammer (2014), where the treatment variable z_i corresponds to the curb weight of the vehicle.

Illustrating the Bias from a Lack of Randomness in an Econometric Context. In its simplest form, researchers implement estimation of a severity model using disaggregate data that is given by:

$$S_{kn} = \beta'_k x_{kn} + \varepsilon_{kn} , \tag{5}$$

where S_{kn} is an injury-severity function determining the probability of injury severity category k for vehicle occupant n, x_{kn} is a vector of explanatory variables that affect the occupant's injury severity level k, β_k is a vector of estimable parameters, and ε_{kn} is an error term. Because the severity outcomes are clearly described in police accident reports, ranging from a fatality to vehicle damage only, researchers can approach the problem using alternative methodologies to analyze discrete data, including ordered models and mixing models that account for preference heterogeneity.

In this model, the effectiveness of seatbelt use, for example, in reducing a fatality is estimated by including in the specification whether a seatbelt was used when an accident occurred (e.g., Eluru and Bhat, 2007). The police officer investigating an accident will report this variable in the police accident report after inspecting the accident. But, as discussed, the use of a non-random sample will still cause the estimate of the effect seatbelt use to be biased and as noted, a selection equation cannot be used that will be uncorrelated with all the omitted influences caused by selectivity bias that influence the occurrence of an automobile accident. Even a random sample will lead to bias unless one makes the implausible assumption that seatbelt use is random and exogenous.

A minority of researchers also have estimated the determinants of severity without taking an econometric approach by performing simple data comparisons. For example, we noted that Evans (1986) compares the severity outcomes of pairs of passengers in the same car involved in an accident, with one passenger wearing and the other passenger not wearing a seatbelt. Although this approach explicitly controls for differences in vehicle occupants who are involved in different accidents, it will still yield biased estimates of the effectiveness of seatbelts in reducing the probability of a fatality because it is based on a non-random sample of automobile travelers. That is, it consists of only those travelers who travel with companions who have distinctly different habits of wearing a seatbelt than they do. A sample designed to include automobile travelers' distinct seatbelt wearing habits, which are correlated with the travelers' attitudes toward safety, will yield biased estimates because seatbelt use will necessarily be correlated with the driver's attitude toward safety. The finding that motorists who wear seatbelts are less likely to be involved in a fatal accident may simply reflect that safer drivers, who are less likely to be involved in a fatal accident than are other drivers, are more likely to wear seatbelts.

Is Selectivity Bias Fatal? We recognize that several arguments could be made to justify a disaggregate estimation approach under certain circumstances, but all of them can be debunked. First, it could be argued that the identification problem is mitigated if estimates of the effect of seatbelt use, for example, on injury severity, $\frac{\partial \pi_{it}^{j**}}{\partial z_i}$, could be interpreted as proxy estimates of $\frac{\partial \pi_{it}^{j*}}{\partial z_i}$ because they were obtained from empirical models that are insensitive to the inclusion of additional control variables or because they are based on plausibly exogenous instrumental variables for seatbelt use z_i.

However, this argument obscures but does not address the fundamental identification issue. Note that differentiation of equation (2) yields:

$$\frac{\partial \pi_{it}^{j**}}{\partial z_i} = \frac{1}{\pi_{it}^{*}}\left(\frac{\partial \pi_{it}^{j*}}{\partial z_i} - \pi_{it}^{j**}\,\frac{\partial \pi_{it}^{*}}{\partial z_i}\right). \qquad (6)$$

Even though π_{it}^{*} is observable, equation (5) implies that $\frac{\partial \pi_{it}^{j*}}{\partial z_i}$ can be recovered from an estimate of $\frac{\partial \pi_{it}^{j**}}{\partial z_i}$ only if $\frac{\partial \pi_{it}^{j*}}{\partial z_i} = 0$, which is the same identifying assumption indicated above. This relationship is implausible because it is difficult to believe that *any* vehicle attribute or driver/roadway attribute would affect the unconditional probability of getting in a severe accident without affecting the probability of getting in any accident. Indeed, our theoretical framework shows that, in general, drivers' decisions will influence the (unconditional) probabilities of getting in accidents of all types of severity.

Second, our framework, which stresses that drivers make many endogenous decisions before and during their trip, reveals that disaggregate approaches cannot be used to obtain consistent estimates of the determinants of the *marginal* probability π_{it}^{*}. This is a critical limitation for two reasons. First, explaining the probability of getting in any kind of accident is one of the most important objects of interest to policymakers. Second, it is not possible to use the marginal probability of getting in an accident as a selection equation to obtain consistent estimates of the determinants of accident severity, which do not suffer from selectivity bias.[9]

Third, it could be argued that an estimate of the *conditional* effect of any determinant of safety on reducing severe or fatal accidents (conditional on any accident occurring) is informative in its own right. But the estimate will still suffer from endogeneity bias that cannot be addressed using disaggregated data. To see this in the case of seatbelts, suppose there was some confounding cause of accidents that was unobservable and correlated with seatbelt use, such as whether the driver was extremely distracted and delayed fastening her seatbelt. Then researchers would need to block the pathways from this confounding variable to both π_{it}^{j**} and π_{it}^{*}. That is, by not being able to study π_{it}^{j*} directly, researchers would need to make an additional identifying assumption. For example, the assumption would require that being extremely distracted does not affect the likelihood of getting in a fatal accident, even if it affected the likelihood of getting in any type of accident. This is not only an implausible assumption, but it appears that researchers taking a disaggregate approach to estimate conditional probabilities may not even be aware that they are making it.

Fourth, it could be argued that standard econometric corrections, such as selection adjustment models or Instrumental Variables (IV), could mitigate the biases that are inherent in disaggregated models. We contend that these methods are analytically futile in this context because a credible exclusion restriction, which is required for these methods to be effective, is virtually non-existent. For example, Xu et al. (2017) utilize a Heckman selection model to analyze signalized intersections, treating roadway segments with zero crashes as censored observations. However, such an approach still fails to satisfy the exclusion restriction required for structural identification. Any variable that influences the selection into the crash dataset—such as the physical configuration of the intersection or localized traffic volume—is inextricably linked to the unobserved risk-taking behavior of the drivers navigating those segments. In the absence of an exogenous instrument that influences crash probability without affecting crash severity, these corrections merely exchange one form of specification bias for another.

---

[9] Eluru and Bhat (2010) effectively take a selectivity approach by jointly modelling seatbelt use and accident severity. In this approach, the endogenous treatment of seatbelt use takes the role of a selection equation to reduce the biased parameter estimate of seatbelt use in the accident severity equation. But the authors do not have clean variation in seatbelt use that is uncorrelated to the determinants of accident severity. Thus, identification is achieved by the choice of functional form, which does not address the fundamental endogeneity problem in their model.

The search for a credible IV in the safety literature remains a largely unfulfilled endeavor. Giles (2001) highlights the technical application of Heckman's methodology to road crash costs, yet even in sophisticated applications, the results remain highly sensitive to the chosen exclusion restriction. In the context of driver behavior, an instrument must influence the adoption of a safety device (e.g., seatbelt use) while remaining orthogonal to the driver's latent preference for risk. Because safety-conscious drivers self-select into safer vehicle configurations and more cautious driving behaviors simultaneously, the omitted variable of driver temperament contaminates nearly all candidate instruments. Without a variable that can bypass this behavioral feedback loop, the IV approach in disaggregated safety research remains an exercise in statistical modeling rather than causal discovery.

Finally, the defense that disaggregated studies remain informative for engineering design or trauma care is, in a policy context, an empty vessel. These fields treat the crash event as a static, exogenous given. While biomechanical data may measure a vehicle's technical potential for protection, it cannot predict its real-world safety performance when human risk-taking is endogenous. To rely on crash-conditioned data to guide safety policy is to mistake the post-hoc anatomy of a collision for the causal economics of the accident

The Aggregate Approach

Independently of research based on the controlled environment and disaggregate approaches, researchers have conducted safety studies that rely on observational data and allow for the analysis of unconditional accident probabilities by collecting aggregated data that includes drivers who did and did not get involved in accidents. The aggregate approach attempts to identify a causal relationship between the fatality rate per vehicle mile traveled (VMT) at the national, state, or regional level over a given time period as a function of safety policy variables, such as speed limits and seatbelt laws, and other influences, such as alcohol and drug consumption. Because the approach aggregates the outcomes of all drivers, it includes the many drivers who never got into a fatal accident by construction.

In the context of our model of driver behavior, if data were collected on the universe of all vehicles on the road for a given time period along with the severity outcomes in accident reports for those vehicles involved in an accident, measures of an accident or fatality rate could be constructed that are analogous to π_{it}^{j*}. Thus, for example, the effect of seatbelt use on the fatal accident rate, $\frac{\partial \pi_{it}^{j*}}{\partial z_i}$, could be identified provided we had variation in z_i that is orthogonal to other elements of X_{jt}.

The limitation of the aggregate approach is that it restricts the questions that can be asked about how to improve automobile safety. For example, aggregated data may not be available for particular socioeconomic groups of drivers, such as teenagers or less-affluent motorists, who are more likely to get into accidents than other groups of drivers. Thus, it may not be possible to estimate the effect of the introduction of states' seatbelt laws on the fatality rate of teenagers and less-affluent households, which may be of particular interest. So, researchers may find that seatbelt laws reduce fatality rates, but they can only speculate about the primary sources of the safety improvement and cannot use the findings to target safety policies more effectively.

Researchers have made effective use of the aggregate approach to estimate broad impacts of policy changes. For example, Dee (1998) and Cohen and Einav (2003) leveraged the staggered rollout of state-level mandatory seat belt use laws to identify the effects of seatbelt use on the rate of overall driving fatalities. In those studies, the change in seat belt use is constant with aggregation (state-year combination), and the entire universe of fatal accidents is reported for each state. Thus, these studies provide consistent estimates of the effects of seatbelt use on the rate of overall driving fatalities. Anderson, Liang, and Sabia (2024) updated Cohen and Einav's study by incorporating twenty-two additional years into the analysis and by applying a new econometric estimator. They obtained estimates of the effects of state-level mandatory seat belt laws on overall fatalities, which were consistent with Cohen and Einav's estimates.

5. Comparing the Approaches' Findings and Cautions About Policy Evaluation and Efficacy

Table 1 summarizes the findings from a selection of studies taking different methodological approaches to estimate the extent that wearing seatbelts reduces driving fatalities. We acknowledge that these studies estimate different estimands across varying contexts. However, the contrast in the findings, which cannot be explained by the variation in the context of the studies, serves as a stark illustration of the upward bias inherent in methodologies that ignore behavioral offsets and that condition on the crash.

Interestingly, the studies that take the controlled environment and disaggregate approaches, which we argued are particularly susceptible to bias that could inflate the safety effects of seatbelts, find that seatbelts produce very large reductions in auto fatalities on the order of 40% to 60%. In contrast, studies that take the aggregate approach, which we argued are not subject to the same bias that affects estimates obtained from the controlled environment and disaggregate approaches, find that wearing seatbelts produce notably smaller reductions in auto fatalities on the order of 10%.

Circumstantial evidence on seatbelt use and automobile fatalities in the United States in recent decades suggests that the smaller estimates obtained from the aggregate approach are more plausible than the larger estimates obtained from the alternative approaches. As shown in figure 1, highway fatalities have declined more slowly during the 2000s than in previous decades, roughly 2% from 2000 to 2023. During the same period, because of stronger and more comprehensive seatbelt laws at the state level, greater enforcement of those laws, and public awareness campaigns, seatbelt use in the United States has increased from roughly 70% in 2000 to roughly 92% in 2023.[10] Thus, the 30% increase in seat belt use during the period is associated with a 2% decrease in auto fatalities, or an elasticity of roughly 7%, which is much closer to the estimates obtained from the aggregate studies than it is to the estimates obtained from the controlled environment and the disaggregate studies. Of course, this comparison does not hold other influences on automobile fatalities constant. But it is difficult to identify other changes in drivers and the driving environment during that period that could have significantly reduced the effect of the increase in seatbelt use on fatalities.

This exercise offers two cautions when researchers seek to offer government safety policy recommendations. First, as we have argued, the benefits of government policies that mandate seat belt use will be biased upward if they are based on the controlled environment and disaggregate approaches. Second, it is important that the costs of safety policies, such as mandatory seatbelt laws, are not overlooked. For example, Thaler and Rosen (1976) and Mannering and Winston (1987) found that although federal law in 1968 required seat belts to be installed in all vehicles except buses, many motorists eschewed their safety benefits based on a rational cost-benefit assessment of the time and bother costs to fasten seat belts and their effect on reducing the probability of a fatal accident.

---

[10] These figures are from the National Highway Traffic Safety Administration's (NHTSA) National Occupant Protection Use Survey (NOPUS).

Even by 1985, when New York was the first state to introduce a mandatory seatbelt law, seatbelt use in the nation was only 19%. By 1995, when 49 states had introduced some type of mandatory seatbelt law (New Hampshire has yet to introduce one), seatbelt use in the nation was roughly 68%. Thus, roughly one-third of US motorists found that the disutility of wearing a seatbelt was sufficiently onerous that they were willing to disobey the law and eschew the safety benefits of wearing a seatbelt. Undoubtedly, during the period when seatbelt laws were being introduced by the states, a large share of motorists could have concluded that they incurred time and bother costs from being forced to use seatbelts that exceeded the benefits that they perceived from wearing one. The unfavorable cost-benefit comparison would have been modestly exacerbated by the increase in the vehicle purchase price that motorists would have to pay for their vehicle to be equipped with seatbelts.

Of course, seatbelt use is much higher today and there is little evidence that a notable share of motorists is, on net, incurring costs from using them. But between 1985 and 1995, well-intentioned policymakers, who believed that seatbelt use would reduce the probability of a fatality by 40% to 60%, could have been influenced to prematurely introduce mandatory seatbelt laws even though they were opposed by nearly two-thirds of the public in a 1984 Gallup Poll.[11] Those laws would have been ill-advised because they produced fewer benefits than were expected and were exceeded in many cases by motorists' costs from being required by law to wear them. It also appears that policymakers prematurely mandated in 1998 that automakers install airbags in all new cars and light trucks despite consumers steadily adopting them and automakers installing them in a manner that was consistent with cost-benefit analysis.[12]

The cautions we have made about the inflated benefits and neglected costs to motorists from government mandates of vehicle occupant safety devices cast strong doubt on Viano's (2024) assertion that, in general, government automobile safety regulations requiring vehicle occupant safety devices and automobile safety features have been cost effective because the evidence in government reports and transportation journal articles that Viano references is compromised by

---

[11] https://tpmblegal.com/how-seatbelt-use-has-changed/

[12] Mannering and Winston (1995) found that, on average, motorists were willing to pay the average cost of installing air bags in their vehicles and that automakers were steadily installing airbags on those vehicles for which motorists were willing to pay the average cost of air bag installation. Nonetheless, in 1998, federal law required that all cars and light trucks sold in the United States have air bags on both sides of the front seat without policymakers carefully assessing whether such a requirement was justified on cost-benefit grounds, accounting for the welfare loss to motorists who valued air bags at less than the cost that was passed on to them through higher vehicle prices.

the problems identified here. First, none of the authors identified and discussed the potential for selectivity bias and how they addressed it to prevent their estimates of safety benefits from being inflated. Second, none of the authors included cost estimates for those motorists who were forced to pay higher vehicle purchase prices for automobile safety devices and features that they valued at less than those higher costs. Without complete and credible causal evidence, Viano's assertion simply distorts the efficacy of government safety regulations.

6. Why Have Transportation Researchers Continued to Use the Disaggregated Approach?

Notwithstanding the significant shortcomings of the disaggregated approach to estimating the causes of accidents that we have discussed here, a large body of transportation research continues to estimate disaggregated models of the determinants of highway safety. We contend that this practice continues because the findings are not used to recommend new automobile safety policies and are therefore not subjected to widespread debate that would eventually shine a light on the flawed methodology of disaggregated models that produced the findings. At the same time, transportation engineers have an interest in using the findings to potentially guide construction decisions.

Lack of Policy Relevance

Transportation researchers would be more aware that the disaggregated approach to analyzing automobile safety is not useful and would be more likely to abandon it if the ultimate goal of this research program were to produce scholarly research that could potentially inform policymakers' efforts to improve highway safety. If that were the goal, then the credibility of researchers methods and findings would be scrutinized more carefully and debated in the context of cost-benefit analyses that seek to provide policy guidance.

However, this branch of safety research has not been taken seriously by government policymakers and has not led to any innovative and effective automobile safety policy recommendations. One can point to states' mandatory seatbelt laws as an example of an automobile safety policy that eventually turned out to be effective. But those laws were hardly spurred by the findings of disaggregated models of the determinants of automobile accidents. As the scholarly research program currently stands, it effectively consists of demonstration papers that use different econometric methods and data sets to obtain parameter estimates, but that do not reach any substantive conclusions that have accumulated and can guide new safety policies or policy reforms.

Potential Use in Highway Engineering Projects

Although the findings from disaggregated models of automobile safety have not been useful for government safety policy, interest exists in using them for practical highway engineering applications at the project level. The design of U.S. highways is based on standards outlined in official federal documents, particularly those maintained by the Federal Highway Administration as codified in the *Federal Lands Highway Manual* and the *Code of Federal Regulations*, specifically 23 CFR Part 625 – Design Standards for Highways.

In theory, those standards could be accurately guided by the findings from econometric work from which it were possible to determine design features that optimize safety. For example, a disaggregated analysis of the determinants of accidents that includes detailed characteristics of the roadway in the specification could be used to determine the net social benefits from using 6-foot roadway shoulders on a specific highway segment instead of using 4-foot shoulders, accounting for accident, construction, and maintenance costs. Although identification problems compromise the estimates obtained from disaggregated models, some highway engineers may still want to use them to the extent that they believe that even flawed estimates are still preferable to no quantitative estimates.

Historically, this attitude has been harmful because highway engineers' use of flawed empirical parameters has compromised highway design and led to considerable waste. Small and Winston (1988) critique the estimate of the relationship between pavement life and thickness that was determined as part of major road test carried out by the American Association of State Highway Officials (AASHO) between 1958 and 1960. The empirical results, published in Highway Research Board (1962), were incorporated into the standard pavement design guide (American Association of State Highway and Transportation Officials, 1981, pp. 59-62, 102-106) on which most states base their design practice.

Small and Winston identified serious flaws with the empirical work that was used to obtain the results and showed that correcting those flaws implied far shorter pavement lifetimes for thick pavements that were used on most interstate highways. In other words, highway durability was underbuilt. The design flaw that was costing the public billions of dollars in additional annual maintenance costs could be addressed by efficient pavement wear pricing for trucks and optimal investment in durability (Small, Winston, and Evans (1989)). We also should point out that the flawed transportation engineering study that policymakers relied on ignored the perverse behavioral incentives of the fuel tax on truckers' use of the road, which encouraged truckers to use trucks with fewer axles to get better fuel economy and reduce their tax burden but increased the pavement damage that their vehicles were inflicting.

We are not aware of retrospective assessments of the costs of flawed highway designs on safety that have been based on biased empirical results obtained from disaggregated models of the determinants of accidents and their severity. However, given the potentially large magnitude of those costs, engineers should pay more attention to the limitations of disaggregated models before using them in practice.

7. Advances in Computation and Data to Potentially Improve Automobile Safety Research

Advances in computation and the availability of new data sets offer the potential for researchers to make progress in overcoming the fundamental challenges of empirical automobile safety research. Although it is premature to claim that the advances are a silver bullet that immediately address the weaknesses in current safety research, careful use of the advances in the context of the controlled environment, disaggregate, and aggregate empirical approaches to analyzing safety in combination with close attention to the identification issues we have raised here have the potential to yield constructive insights to help improve the analysis of automobile safety.

The Controlled Environment Approach

The major concern with the controlled environment approach is that although researchers can create environments to analyze safety performance that hold constant external confounders (e.g., weather), they cannot account for the fact that the same external confounders may affect the actual driving decisions of motorists (e.g., riskier drivers may be more likely to travel in adverse weather). Thus, experiments of safety performance in controlled environments may be compromised by failing to construct accurate counterfactuals in actual driving environments.

As a potential step forward, consider that the driving algorithms that are used by, for example, Waymo, to navigate autonomous vehicles have been trained successfully on massive quantities of sensory data that are collected from actual driving environments on US roads. Thus, instead of designing controlled physical environments to test automobile safety equipment, researchers might piggyback off the work by firms and research groups to improve autonomous vehicles by designing controlled virtual driving environments to test marginal improvements in safety equipment *in silico*. In the process, researchers would be better able to understand how safety equipment could change the actual driving environment.

In terms of applications, the controlled environment approach *in silico* might be quite successful at, for example, estimating the safety impacts of an automatic emergency braking technology that was modestly more responsive than currently deployed braking technologies because the behavior of those technologies on the road is likely to be well understood by current autonomous driving models. However, it could be much more difficult for the virtual controlled environment approach to estimate the safety effects of a fundamentally new braking technology, where, for example, vehicles communicate directly with traffic signals, because the concomitant changes in driving behavior would be "out of sample" relative to the experience of current training data sets.

The Disaggregate Approach

The weakness of the disaggregate approach is that it relies on police accident reports, enabling researchers to identify only the effects of technologies on safety conditional on a crash occurring. The proliferation of driving sensors on roadways, traffic signals, vehicles, and other physical infrastructure along with widely deployed innovations in video capture and parsing may make progress in addressing this limitation by generating data on driver behavior at *all* times, not just those times when a crash occurs. Collecting these data for motorists when safety devices are and are not available on their vehicles would allow researchers to estimate the effects of safety improvements under relevant counterfactuals. Of course, drivers' offsetting and selection behavior would still be relevant issues, but they may be modeled more easily with comprehensive data and as estimation techniques to exploit "big data" are honed.

The Aggregate Approach

Although the aggregate approach is not subject to the selection issues that plague much of the current auto safety literature, it cannot be used to estimate the heterogeneous effects of a safety technology across different driver and vehicle types and the interactive effects of multiple safety technologies deployed in tandem. Improvements in data collection might enable the aggregate approach to be used for a wider range of questions. For instance, Maheshri, Winston and Wu (2025) used new natural language processing and string-matching techniques to construct a panel dataset of the universe of *all* vehicles and accidents at the calendar year-model year-make-model-trim level. This finer degree of aggregation allowed for the analysis of a previously unobserved (to the researcher) natural experiment; namely, the staggered timing of the rollout of advanced driving assistance technologies (ADAS) to otherwise identical vehicles of different trim levels.

As the degree of detail in administrative datasets, crash reports, and manufacturer data continues to increase, new computational techniques may harness this informational surplus to gain more credible insights into current and new automobile safety questions. Researchers, however, must still be vigilant about identification issues that may be raised by the use of these data.

8. Conclusions as We Enter an Era of Vehicle Autonomy

This paper offers three important conclusions about automobile safety research and safety performance and policy as we enter an era of vehicle autonomy. First, the flaws inherent to conventional disaggregate and controlled environment approaches prevent this line of research from obtaining causal estimates of the determinants of automobile accidents that are useful for safety researchers and policymakers.[13] Causal estimates of the determinants of automobile accidents obtained from an aggregate approach can be identified in limited contexts. Unfortunately, the disaggregate approach has become the standard approach in safety research and as have we have discussed, researchers have continued to use it because the method and the findings have not been assessed on the grounds of whether they are useful in practice, while they continue to be used for applications by transportation engineers.

Second, notwithstanding the limited contributions of safety research, advances in automobile technology and public investments in infrastructure have enabled automobile safety to steadily improve for roughly a century. However, government policymakers have periodically overreacted to occupant safety improvements by prematurely mandating that motorists should use them and that automakers should make their use possible by installing them in all their new vehicles. Those mandates have imposed costs on consumers whose value of the increased safety

---

[13] The same conclusion can be reached about studies that take a disaggregate or a controlled environment approach to explain the determinants of automobile accidents involving pedestrians. Selectivity bias can arise in both the type of motorists and pedestrians who are involved in those accidents.

is less than the costs of time and bother to use the new occupant safety features and the increase in prices they must pay to cover installation costs.

Finally, the most promising source of a significant safety improvement in the future is the major technological advance represented by the widespread adoption of autonomous vehicles. AVs would replace the drivers' optimization problem that we have formalized here with the network optimization problem of determining vehicles' speeds and routings without the threats to safety created by drivers' heterogeneous preferences for risky behavior (Winston and Karpliow (2020), Winston, Yan, Associates (2024)). Initial safety comparisons between autonomous vehicles, based on Waymo's performance, and non-autonomous vehicles find that Waymo's crash rate is much lower than human drivers' crash rate (Kusano et al. (2025)) and that Waymo's automobile liability insurance claims were considerably lower than human-driven vehicles' liability claims (Di Lillo et al. (2024)).[14]

Although AVs' technology is currently being perfected and tested, their widespread adoption is still decades away. Thus, as it was important for policymakers to respond appropriately to the introduction of non-autonomous vehicle safety features, it is important for policymakers to respond appropriately to the gradual introduction of AV technologies by drawing on credible evidence of their costs and benefits, potentially obtained by constructive use of improvements in computation and data discussed in the previous section.

Unfortunately, the recent federal response to the advent of advanced driver-assistance systems (ADAS) reflects a persistent history of prematurely mandating occupant safety technologies based on insufficient causal evidence. This trend is exemplified by the National Highway Traffic Safety Administration's (2024) mandate (FMVSS No. 127), which requires all new passenger vehicles and light trucks to be equipped with automatic emergency braking (AEB)—a key component of ADAS—by 2029.

While identified aggregate models (Maheshri, Winston, and Wu, 2025) find that ADAS technologies have reduced the risk of all-type accidents by 11 to 14 percent and single-vehicle fatal accidents by roughly one-third, federal policy appears to be driven by significantly more optimistic and likely misleading estimates. For instance, Haus, Sherony, and Gabler (2019) utilized a

---

[14] The validity of these comparisons could be challenged if safer drivers self-selected to use Waymo's vehicles and less-safe drivers self-selected to be human drivers. However, such self-selection, even if it has occurred to some extent, is highly unlikely to account for the difference between the safety of autonomous and human driven vehicles.

controlled environment simulation to estimate that AEB could reduce pedestrian fatality risk by as much as 85%. This massive discrepancy clearly illustrates the upward bias produced by simulations that fail to account for behavioral responses to new safety technologies and the complexities of the highway environment.

While the internal deliberations of federal regulators are not public, the official Regulatory Impact Analysis (FRIA) for FMVSS No. 127 demonstrates that the mandate's justification is built upon the very methodological pillars we critique. Specifically, the agency relies on disaggregated statistical evaluations of AEB effectiveness, such as Cicchino (2017), which utilize police-reported crash data to compare involvement rates between vehicles with and without the technology. As we have argued, such comparisons are fundamentally compromised by selection bias on the dependent variable—the occurrence of the crash itself—and ignore the behavioral equilibrium that determines real-world safety performance.

This premature AEB mandate risks repeating two historical errors. First, it may force the widespread adoption of a technology before its long-run, equilibrium impact on societal safety is understood. Second, it imposes a potential welfare loss on consumers. As with the premature mandates for seatbelts and airbags—which, as noted, incurred costs for many—consumers who do not value the actual safety gains of AEB as much as its installation cost (roughly \$500 to \$2,000 depending on the vehicle) will incur a net economic loss.

Although many automobile safety researchers have historically avoided policy debates, the concerns we have raised extend well beyond the academic community. The official regulatory record (NHTSA, 2024) reveals that current federal mandates are being justified by the very track-test and crash-conditioned methodologies that we have argued to be fatally flawed. As we enter the autonomous era, the research community faces a choice of continuing to rely on biomechanical capacity as a proxy for safety, or embracing the causal rigor necessary to optimize real-world safety performance without ignoring human behavioral offsets.

References

American Association of State Highway and Transportation Officials, AASHTO. 1981. Interim Guide for Design of Pavement Structures 1972, Chapter III Revised, 1981, Washington, DC, AASHTO.

Anderson, D. Mark, Yang Liang, and Joseph J. Sabia. 2024. "Mandatory Seatbelt Laws and Traffic Fatalities: A Reassessment," *Journal of Applied Econometrics,* vol. 39, no. 3, pp. 513-521.

Anderson, Michael L., and Maximilian Auffhammer. 2014. "Pounds That Kill: The External Costs of Vehicle Weight," *Review of Economic Studies,* vol. 81, no. 2, pp. 535-571.

Cicchino, Jessica B. 2017. "Effectiveness of Forward Collision Warning and Autonomous Emergency Braking Systems in Reducing Front-to-Rear Crash Rates," *Accident Analysis & Prevention*, 99, pp. 142-152.

Cohen, Alma, and Liran Einav. 2003. "The Effects of Mandatory Seat Belt Laws on Driving Behavior and Traffic Fatalities." *Review of Economics and Statistics*, vol. 85, no. 4, pp. 828–843.

Dee, Thomas S. 1998. "Reconsidering the Effects of Seat Belt Laws and Their Enforcement Status." *Accident Analysis & Prevention*, vol. 30, no. 1, pp. 1–10.

Di Lillo, Luigi, Tilia Gode, Xilin Zhou, John Scanlon, Ruoshu Chen, and Trent Victor. 2024. "Do Autonomous Vehicles Outperform Latest-Generation Human-Driven Vehicles? A Comparison to Waymo's Auto Liability Insurance Claims at 25 Million Miles," https://storage.googleapis.com/waymo-uploads/files/documents/safety/Comparison%20of%20Waymo%20and%20Human-Driven%20Vehicles%20at%2025M%20miles.pdf

Eluru, Naveen, and Chandra R. Bhat. 2007. "A Joint Econometric Analysis of Seat Belt Use and Crash-Related Injury Severity," *Accident Analysis and Prevention*, vol. 39, no. 5, pp. 1037–1049.

Evans, Leonard. "The Effectiveness of Safety Belts in Preventing Fatalities." *Accident Analysis and Prevention*, vol. 18, no. 3, 1986, pp. 229–241.

Giles, Margaret J. 2001. "Heckman's Methodology for Correcting Selectivity Bias: An Application to Road Crash Costs, *SSRN Electronic Journal*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=288275

Haight, Frank A. 1970. "A Crude Method for Estimating Air Travel Exposure in Deaths per Passenger Mile, *Business Horizons*, 13(3), 84–86.

Haus, Samantha H., and Rini Sherony, and Hampton C. Gabler. 2019. "Estimated Benefit of Automated Emergency Braking Systems for Vehicle–Pedestrian Crashes in the United States," *Traffic Injury Prevention,* vol. 20, No. S1, pp. S171-S176.

Highway Research Board. 1962. The AASHO Road Test, Report 5: Pavement Research, Special Report No. 61 E.

Kahane, C. J. 2015. Lives saved by vehicle safety technologies and associated Federal Motor Vehicle Safety Standards, 1960 to 2012 – Passenger cars and LTVs – With reviews of 26 FMVSS and the effectiveness of their associated safety technologies in reducing fatalities, injuries, and crashes. (Report No. DOT HS 812 069). Washington, DC: National Highway Traffic Safety Administration, January.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), 263–291.

Kusano, Kristofer D., John M. Scanlon, Yin-Hsiu Chen, Timothy L. McMurry, Tilia Gode, and Trent Victor. 2025. "Comparison of Waymo Rider-Only Crash Rates by Crash Type to Human Benchmarks at 56.7 Million Miles," *Traffic Injury Prevention,* 1-13. *https://doi.org/10.1080/15389588.2025.2499887*

Maheshri, Vikram, Clifford Winston, and Yidi Wu. 2026. " AI at the Wheel: The Effectiveness of Advanced Driver-Assistance Systems," *Journal of Law and Economics,* May, forthcoming.

Mannering, Fred, and Clifford Winston. 1987. "Recent Automobile Occupant Safety Proposals," in Clifford Winston and Associates, *Blind Intersection? Policy and the Automobile Industry,* Brookings Institution, pp. 68-88.

Mannering, Fred, and Clifford Winston. 1995. "Automobile Air Bags in the 1990s: Market Failure or Market Efficiency?," *Journal of Law and Economics,* volume 38, October, pp. 265-79.

National Highway Traffic Safety Administration (NHTSA). 2024. *Final Regulatory Impact Analysis: Federal Motor Vehicle Safety Standard No. 127; Light Vehicle Automatic Emergency Braking (AEB); AEB Test Devices.* Docket No. NHTSA-2023-0021. Washington, D.C.: Office of Regulatory Analysis and Evaluation, National Center for Statistics and Analysis, April.

Peltzman, Sam. 1975. "The Effects of Automobile Safety Regulation," *Journal of Political Economy*, vol. 83, no. 4, pp. 677–725.

Small, Kenneth A. and Clifford Winston. 1988. "Optimal Highway Durability," *American Economic Review,* volume 78, June, pp. 560-569.

Small, Kenneth A., Clifford Winston, and Carol A. Evans. 1989. *Road Work: A New Highway Pricing and Investment Policy,* Brookings Institution.

Thaler, Richard, and Sherwin Rosen. 1976. "The Value of Saving a Life: Evidence from the Labor Market," in Nestor E. Terlecky, editor, *Household Production and Consumption,* National Bureau of Economic Research.

Thorpe, J. D. 1964. “Calculating Relative Involvement Rates in Accidents Without Using Exposure Data,” *Australian Road Research*, 2(1), 25–36.

TRIP. 2023. *Addressing America's Traffic Safety Crisis: Examining the Causes of Increasing US Traffic Fatalities and Identifying Solutions to Improve Road User Safety*, Washington, DC, June.

Tscharaktschiew, Stefan. 2020. “Why are Highway Speed Limits Really Justified? An Equilibrium Speed Choice Analysis,” *Transportation Research Part B: Methodological,* vol. 138, August, pp. 317-351.

Vertlib, Shani R., Stav Rosenzweig, Ofir D. Rubin, and Aviv Steren. 2023. “Are Car Safety Systems Associated with More Speeding Violations? Evidence from Police Records in Israel,” PLoS ONE 18(8): e0286622. https://doi.org/10.1371/journal. pone.0286622

Viano, David C. 2024. “Government Regulations Have Improved Automotive Safety,” *Traffic Injury Prevention,* volume 25, issue 2, pp. 101-102.

Wilde, Gerald J.S. 1982. “The Theory of Risk Homeostasis: Implications for Safety and Health,” *Risk Analysis,* vol.2, issue 4, pp. 209-225.

Winston, Clifford, and Quentin Karpilow. 2020. *Autonomous Vehicles*: *The Road to Economic Growth?* Brookings Institution Press.

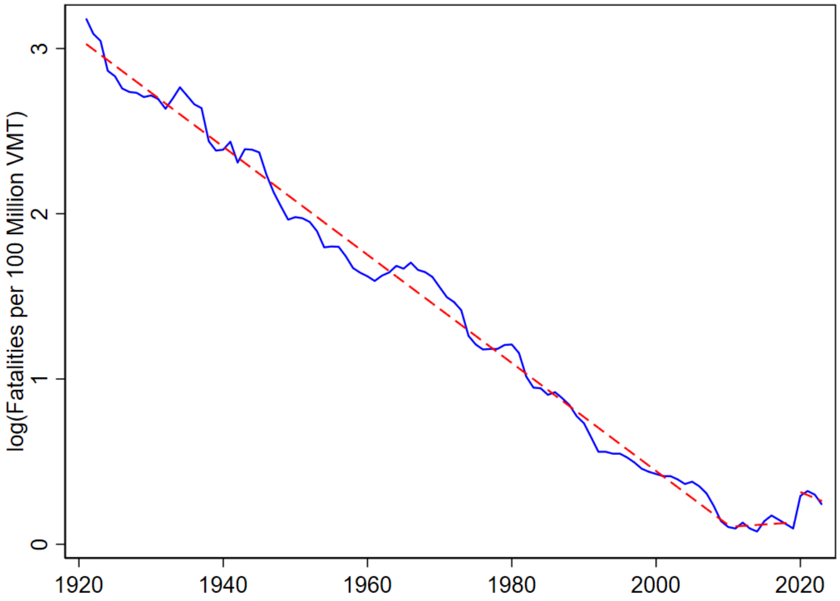
Winston, Clifford, Vikram Maheshri, and Fred Mannering. 2006. “An Exploration of the Offset Hypothesis Using Disaggregate Data: The Case of Airbags and Antilock Brakes,” *Journal of Risk and Uncertainty,* volume 32, March, pp. 83-99.

Winston, Clifford, Jia Yan, and Associates. 2024. *Revitalizing a Nation: Competition and Innovation in the US Transportation System,* Brookings Institution Press.

Xu, X., Wong, S. C., Zhu, F., Pei, X., Huang, H., & Liu, Y. 2017. “A Heckman Selection Model for the Safety Analysis of Signalized Intersections, *PLOS ONE*, *12*(7), pp. 1-16.

Yang, Hai, Hongbo Ye, Xinwei Li, and Bingqing Zhao. 2015. “Speed Limits, Speed Selection and Network Equilibrium,” *Transportation Research Part C: Emerging Technologies,* vol. 51, February, pp. 260-273.

Figure 1. US Automotive Fatality Rate Over Time (in Logs)[a]



[a] Fatalities data from the US Department of Transportation.

Table 1. Illustrative Comparison of Seatbelt Effectiveness Estimates by Methodological Approach

Year	**Authors**	**Findings**	**Notes**
Controlled Environment			
1970	Lave and Weber	40-50% fatality reduction	Use biomechanical evidence from government crash tests.
2015	Kahane (NHTSA)	25-69% fatality reduction	Use biomechanical evidence from government crash tests.
Disaggregate			
1986	Evans	42% fatality reduction	Compares pairs of passengers in the same car, one belted, one unbelted.
2007	Eluru and Bhat	64% fatality reduction	Joint model of seat belt use and accident severity conditional on a collision.
Aggregate			
1975	Peltzman	0% overall, accounting for pedestrian deaths.	Compares trends before and after 1968 federal safety regulations.
1998	Dee	5-6% fatality reduction	Exploits staggered rollout of mandatory seat belt laws by states in diff-in-diff estimation.
2003	Cohen & Einav	4-6% fatality reduction	Exploits staggered rollout of mandatory seat belt laws by states as IV for reported usage.
2024	Anderson, Liang, & Sabia	5-9% fatality reduction	Replicates and extends the Cohen and Einav study.