POLITICS, POLICY ANALYSIS, AND THE PASSAGE OF
THE NATIONAL MINIMUM DRINKING AGE ACT OF 1984

Darren Grant
Sam Houston State University

September 2011

Abstract:
The National Minimum Drinking Age Act of 1984 exemplifies high-stakes legislation
that attracted the interest of the public, legislators, academics, policy advocates, and
executive agencies. This paper explores how these actors combined to generate
intellectual support for this act within the legislative process. Limitations of the
contemporaneous research required that the available evidence be evaluated judiciously.
This did not happen, because it is not fostered by the adversarial nature of the process and
because its most influential participants, executive agencies heavily involved in traffic
safety, lacked the necessary neutrality and expertise.
POLITICS, POLICY ANALYSIS, AND THE PASSAGE OF THE NATIONAL MINIMUM DRINKING AGE ACT OF 1984

Darren Grant
Department of Economics and International Business
Sam Houston State University
Huntsville, TX 77341-2118
dgrant@shsu.edu

Abstract: The National Minimum Drinking Age Act of 1984 exemplifies high-stakes legislation that attracted the interest of the public, legislators, academics, policy advocates, and executive agencies. This paper explores how these actors combined to generate intellectual support for this act within the legislative process. Limitations of the contemporaneous research required that the available evidence be evaluated judiciously. This did not happen, because it is not fostered by the adversarial nature of the process and because its most influential participants, executive agencies heavily involved in traffic safety, lacked the necessary neutrality and expertise.

JEL Codes: I18, K14, N42

Keywords: drunk driving; traffic safety legislation; policy analysis

---

1 This research was sponsored by a grant from Choose Responsibility. A companion paper, Grant (2011), documents how academic findings on the effects of three major laws—the minimum legal drinking age, zero tolerance laws, and .08 per se laws—become much less favorable over time and explains this evolution in terms of changes in study design and an “early adopter effect.” I appreciate comments from participants at the 2011 Public Choice Society meetings, the research assistance of Kristin Boykin, and the cooperation of several interviewees.
In general the influence of the public, whether directly or through political institutions, has been pernicious to traffic safety. It comes and goes, filling in the troughs between peaks of more exciting events; it seizes on issues without concern for the relevance or tractability of the problems; it proposes “solutions” which are at best naïve and at worst absurd, and above all it demands action even where action may be only a waste of money.

—Frank Haight (1985)

Although it is hard to challenge Haight’s characterization of...the political process in the area of traffic safety, his proposed solution of entrusting the issue to low-profile agencies...seems wishful and unrealistic. His viewpoint neglects the fact that the recognition of any condition as a social problem is a political matter. It is not helpful for underdogs in the political game to pick up their chips, denounce the rules...and look elsewhere, when the political game is the only game in town. Even the experts are forced to play it, often as mere adjutants to parties with less sophistication but greater involvement and determination.

—H. Laurence Ross (1992, p. 174)

The common American ideal of federalism holds that the states, “laboratories of democracy,” experiment with various approaches to solving social problems; the most promising thrive and the others wither. This ideal assumes, or requires, that these alternative approaches can be evaluated accurately. Researchers readily recognize that this need not be so. Even when using appropriate evaluation methodologies, the relative and absolute merits of any given policy are often uncertain, because of differences across studies in research design, data, and the interpretation of the evidence.

The academic community deals with this uncertainty patiently, amassing studies and counter-studies over decades, and, given sufficient time and study, can often (though not always) arrive at a reasonable degree of consensus. Policymakers, in contrast, rarely have this liberty. Solutions to many problems are needed more urgently, so that the evidence available at the point of decision may be deficient in both quality and quantity. How does the political theatre, in which the evidence is presented, summarized, and evaluated, respond to these deficiencies? Is it
structured to do so well?

General, generic answers to these questions, which typically emphasize the difference between intellectual idealism and political reality, are abundant. Specific, comprehensive, context-rich answers—which are required to truly understand the past or to make realistic prescriptions for the future—are not.² These require precise descriptions of what could have been known, or anticipated, about a policy’s likely effects at the time it was adopted; historical context; and institutional specifics that clarify the relevant characteristics of the political process and of the key actors in that process, and show how they influence the outcome. All of this can be obtained only by detailed study of specific policies or specific institutions.

Accordingly, we conduct a comprehensive narrative of the process leading to the passage, in 1984, of the National Minimum Drinking Age Act (NMDAA), “one of the most thoroughly evaluated social interventions of our time” (Ross, 1992). This act, which provided strong and ultimately successful incentives for all states to raise the minimum legal drinking age (MLDA) to twenty-one, was empirically justified using inferences drawn from the experience of a small number of “laboratory states” that had recently raised their MLDAs. Transcripts of

² This paper appears to be the first extended narrative to comprehensively address the questions posed above. The closest related works, both retrospective in spirit, appear to be Henig (2008), who compares how two forums, academia and the media, resolve differences about research findings in the area of charter schools, and Tanenbaum (2009), who examines how Pay for Performance came to be implemented in Medicare despite little concrete evidence on its effectiveness. While neither focuses on the role of evidence in the policy process, as this study does, similar themes pervade all three papers, suggesting at least some degree of universality.

Two other, related literatures have a future orientation: metapolicy—“policy on how to make policy” (Dror, 1971, p. 74)—and robust political economy, which seeks to know “which institutions perform best when people have limited knowledge and are prone to self-interested behavior” (Pennington, 2011, p.3). Both stand in contrast to traditional political and “decisionist” models of policymaking, which emphasize power or technocratic merit, respectively, in favor of a conceptualization that “link(s) the intellectual and political, economic and motivational aspects of the process” (Majone, 1989, p.148). Neither contains the type of extended analysis conducted here.
congressional subcommittee hearings and other sources show how this evidence was employed within the political arena, while dozens of subsequent empirical studies are used both to more accurately assess the effects of the increased MLDA—to give the benefit of hindsight—and to identify the effect of different study features on their conclusions.

Because the available evidence on the effects of a raised MLDA was deficient in quantity and quality, it should have been evaluated with judiciousness and discernment. In practice, however, these qualities were largely missing, for three reasons that still apply today. First, the adversarial aspect of the political system excluded and drowned out the relatively quiet voices of those who, by temperament or training, possessed such judiciousness. This was true even though the NMDAA was bipartisan and the temper of the debate was civil. Second, the forces for and against the raised MLDA were not evenly matched in technical or political skill, further weakening the efficacy of the adversarial process. Finally, the key government agency that Congress relies upon for traffic safety advice, the National Highway Traffic Safety Administration (NHTSA), exhibited neither the independence required for judiciousness nor the technical skills required for discernment. These skills’ absence was not accidental, but rather was a consequence of NHTSA’s organizational design—in particular, the “make or buy” decision for traffic safety research—which itself supported the agency’s objectives.

Ultimately, for the NMDAA, the political process amplified, rather than dampened, the literature’s overly-optimistic predictions of its likely effects. This is not an isolated occurrence: similarly unrealized optimism was expressed over more recent traffic safety legislation promoted by Congress, in part due to the active support of NHTSA, whose organizational design (for traffic safety research) remains unchanged to the present day.
Section I. A Brief History of the National Minimum Drinking Age Act.

Legislative activity to curtail youth drinking and drunk driving in general began in the late 1970s, a counterreaction to increased permissiveness earlier in the decade, when twenty-nine states lowered their drinking ages. From 1976-1980, thirteen states raised their drinking ages, generally by one year. Between 1981 and 1983, twelve more states raised their drinking ages, thirty-four states adopted per se blood alcohol concentration (BAC) limits, and eleven adopted administrative license revocation. Activity crested between 1984 and 1986, with the passage of hundreds of state laws, as documented in Table 1, based on Howland (1988).

This activity was associated with three concomitant social changes, two of which are also documented in the table. The first was an increased social awareness of the dangers of drunk driving: media coverage of the issue, almost wholly absent during the 1970s, grew rapidly after 1981. Hundreds of stories appeared in major newspapers, and dozens of stories in magazines, during the next quadrennium. Coverage in other media increased as well:

I can see it from my experiences of ten, twelve years ago as Secretary, when if I could get one TV camera to come to a hearing or a meeting about drunk driving, I thought we were very fortunate. They might stay as long as ten minutes. Our first hearing in Oklahoma City [of the Presidential Commission on Drunk Driving] we had four television cameras, twelve radio stations, and eight or nine newspaper people there. Two of the cameras stayed half a day and two stayed all day long. (John Volpe, former Secretary of the Department of Transportation and then-Chairman of the Presidential Commission on Drunk Driving, H1, 1983, p. 273.)

This awareness translated into action: hundreds of organizations were founded whose purpose was to curtail drunk driving. The best known of these, Mothers Against Drunk Driving (MADD), played a key role in lobbying for the NMDAA and subsequent legislation, including

---

3 This quote, like many others to follow, comes from the transcript of a Congressional hearing or Congressional debate. These events are each listed in the chronology in the Appendix, labeled H0-H6, and cited within the text using that appellation.
.08 per se BAC limits, open container laws, and zero tolerance laws, and is still active today.

The final social change, in attitudes, can only be documented qualitatively:

For...decades, the enormous toll of death and injury that occurred in the United States was regarded as accidental in almost a cosmic sense. The statistical toll of road accidents was collected and reported with an air of fatalism similar to attitudes toward earthquakes, tornadoes, or other natural disasters. At the same time, the...paradigm of responsibility began and ended with the personal fault of the parties to the accident. The public perception now in the United States...is that the manner in which ...laws are drafted and enforced can have important effects on highway deaths and injuries (Zimring, 1988).

The American public is far less tolerant of drunk driving that they were ten years ago. It’s no longer funny for Johnny Carson to joke about the issue. (Judith Stone, Director, Federal Affairs, National Safety Council, H5, June 1988, p. 24.)

There really was a cultural shift in the 1980s I believe to where impaired driving was no longer an accepted part of American culture. (Jeffrey Runge, Administrator, NHTSA, H6, 2002, pp. 12, 16.)

While most legislative activity occurred at the state level, the issue also received attention within the federal government, partly from concern about “blood borders” created by youth driving across state lines to take advantage of a lower MLDA. This took several forms, as documented in the extensive chronology of the activity surrounding the passage of the NMDAA that is located in the Appendix. Transportation bills offered financial incentives to the states to adopt various drunk driving countermeasures, including but not limited to higher drinking ages. President Reagan appointed a highly-visible Presidential Commission on Drunk Driving, which held nationwide hearings and ultimately issued dozens of recommendations on the issue. Both NHTSA and the National Transportation Safety Board (NTSB) weighed in with written reports and testimony before Congressional subcommittees, which held several hearings on the problems of drunk driving and teenage drinking.

In 1983 and 1984 these forces built to a fever pitch and provided the impetus for strong federal action. In November, 1983, the Presidential Commission on Drunk Driving, following in
the footsteps of MADD, the National Safety Council (NSC), the Insurance Institute for Highway Safety (IIHS), and the NTSB, formally recommended establishing a national drinking age of twenty-one in order to improve traffic safety. A Gallup Poll taken earlier that year indicated that over three-fourths of the country was in support. Multiple bills or amendments to mandate or encourage the raised drinking age were proffered in Congress. The primary objector, President Reagan, changed his position in June, 1984, and the NMDAA was law one month later. Challenges to the law’s constitutionality were exhausted by 1987; the next year the last two states raised their MLDAs to 21.

Section II. Evidence on the Effects of the Minimum Legal Drinking Age: Then and Now.

The country’s enthusiasm for a raised MLDA was matched by its supporters’ optimism about its expected effect on traffic safety:

Nearly every state that has raised the drinking age to twenty-one has produced a significant drop in the (sic) teenage driving fatalities. In the state of New Jersey...the rate dropped by twenty-six percent; Illinois, it has fallen twenty-three percent; in Michigan, thirty-one percent. (President Ronald Reagan, Remarks on Signing HR 4616 into Law, July 17, 1984.)

To some degree, however, this optimism contrasted with the incompleteness of the evidence on how the drinking age affects traffic safety—the law’s raison d’être. The complete literature through 2009, forty years in length, is illustrated in Figure 1, taken from Grant (2011). (This includes all studies published in an academic book or refereed journal that estimate the effect of the raised MLDA on the affected population.) This “bubble plot” illustrates many study features at once: the horizontal axis represents the publication date, which is generally a couple of years after its data terminate, and the vertical axis represents the estimated percentage effect on
fatalities involving drivers affected by the changed MLDA, with insignificant estimates set to zero. The volume of each bubble represents the number of academic citations in Google Scholar as of June 2009, with a minimum bubble size so that uncited studies are not eliminated. (Later studies have less time in which to be cited, of course.) Bubbles ringed in black circles are supported by external funding, generally from the National Institute on Alcohol Abuse and Alcoholism (NIAAA). Finally, the color of the bubble represents the study design. Blue represents quasi-experimental designs, which generally compare the change in fatalities in one, or a few, law-adopting states with that in control states that do not adopt the law. (Sometimes these changes are trend-adjusted using an ARMA model, and some control groups are age-related.) Purple indicates pooled time-series cross-section (TSCS) regressions, which include control variables but not state and year fixed effects; white indicates panel regressions, which do include these fixed effects; and red represents cross-section regression. Panel designs are preferred: they combine the before-after quality of quasi-experimental analyses with the breadth and explicit inclusion of control variables that are found in pooled TSCS regressions.

The MLDA literature contains studies of lowered drinking ages, based on data from the 1970s, and subsequent studies of raised drinking ages. The two are distinguished by the diagonal line in the figure. One might expect the long term effect of raising the drinking age from 18 to 21 to be equal and opposite that of moving in the reverse direction, unless the lower drinking age develops social conventions that prove difficult to dislodge, in which case it would be smaller. Nevertheless, evidence on the effects of lowered MLDAs was almost wholly absent from the testimony we have reviewed, though the best of this evidence (Cook and Tauchen, 1984, discussed below) spanned more states and years than any raised MLDA study could then muster. This remains a mystery: to our knowledge, no one tried to justify excluding these studies.
For the literature as a whole, as well as its raised-MLDA and lowered-MLDA components, the differences in study findings are tremendous, with the estimated effect spanning more than thirty percentage points. Some of these are relatively unsystematic, stemming from differences in the dependent variable (crashes or fatalities, all accidents or only night accidents, scaling by population or miles travelled, etc.), the states and control groups used in quasi-experimental analyses, and the control variables included in regression analyses.

Other differences are systematic. These pertain to study design and execution date, and stem from an evolutionary process documented in Grant (2011) for this and other drunk driving literatures. Early studies, dominated by quasi-experimental methods, yield highly variable yet generally favorable conclusions. As the number of law-adopting states and post-law years grow, these are supplanted by pooled TSCS and, eventually, panel regressions, which are less variable and much less favorable. This can be seen for studies of the lowered drinking age, in the left part of Figure 1, and is even more striking for studies of the raised drinking age, on the right. There the number of law-adopting states and post law-adoption years quickly becomes large, permitting extensive use of panel methods. The two panel analyses of lowered MLDAs, Cook and Tauchen (1984) and Weinstein (1987), find traffic fatalities among the affected ages increase by six or seven percent. Among panel analyses of raised MLDAs, early studies (whose data end before 1990, and which are discussed below) find estimates of about 13%, but later studies’ estimates average, again, six or seven percent. (The estimates in these later studies—Dee, 1999; Eisenberg, 2003; Young and Likens, 2000; Young and Beilinska-Kwapisz 2006; Polnicki et al., 2007; and Miron and Tetelbaum, 2009—range fairly uniformly from 3-11%.)

The only other MLDA study using modern methods, Dobkin and Carpenter’s (2009) regression discontinuity analysis, finds that motor vehicle fatalities among individuals just shy of their twenty-first birthday are 14% below those of individuals just past that birthday. That estimate,
This evolutionary semi-convergence in method and findings is a natural consequence of the ethos of academia, which is organized as an “independent, collective, cumulative, open-ended enterprise of knowledge creation and testing” (Henig, 2008, p. 232). This ethos embraces diversity in the identities of those investigating a topic and the methods used in doing so, allowing alternative approaches to compete over time. The quality of the analysis is emphasized over the ease or speed of execution. And each researcher is expected to make a good faith effort to find the “correct” answer to the question posed, without omitting contradictory evidence or overstating the case. This process often requires a decade or more to culminate, because academic studies take years to complete and publish (and, in some cases, refute), and many such studies may be needed to ultimately settle the question. These properties were on display for the topic on which Henig focuses, charter schools, as they were for the MLDA. Throughout the 1980s, conflicts arose in the MLDA literature over measurement (Williams et al., 1983 vs. Males, 1986; also Hammond, 1973 vs. Zylman, 1974), study design (Garber, 1988), and execution (General Accounting Office, or GAO, 1987), which are now substantially resolved.

Policymakers, however, are rarely afforded this much time. Accordingly, the limited evidence that is available at the point of decision should be carefully scrutinized and judiciously evaluated. This certainly would have been merited in 1983 and 1984 regarding the raised MLDA. As Figure 1 illustrates, the published (or soon to be published) studies then available were relatively few in number, weak in design, and variable in their findings. Also, the GAO (1987) shows that some did not adhere to generally accepted methodological standards, a problem even more common in those contemporaneous studies that were never published. It

while sound, is local, and cannot be used to infer the percentage reduction in fatalities across all ages affected by raised MLDAs. This is probably much smaller: Dee (1999) and Miron and Tetelbaum (2009) both find that outlawing drinking among eighteen and nineteen year-olds hardly changes fatalities, but outlawing it for twenty year-olds has a substantial effect.
should have been—and, eventually, was—a focus of Congressional hearings to identify the best available evidence on the effects of a changed MLDA and weight it accordingly.

There is, however, another, more subtle sense in which judiciousness is called for, which derives from the fact that early studies are also systematically more favorable than later, more comprehensive analyses with improved study designs. Collectively, then, these early estimates were a biased predictor of the long run, nationwide effects of a raised MLDA. This is not just a matter of hindsight: one could have justifiably been concerned about this at the time, because the technical and conceptual groundwork for these concerns had already been laid.

Technically, the quasi-experimental methodology used in most early studies of this issue (and which continues to be common in the traffic safety literature) has three important limitations. First, these studies operate on short time scales. Often, the number of post-law years analyzed is just one or two; the number of pre-law years may not be much larger. Without direct controls for economic or drinking-related factors, which are rarely present, it can be difficult to accurately identify baseline short run and long run trends from which to extract the effect of the law. Second, these studies assume that the paired law-changing and control states are otherwise equivalent. This can bias estimates, because this assumption is not empirically supported (Grant, 2010), and does overstate statistical significance, because state-specific random effects are neither allowed in the underlying empirical model nor identified. (This method assumes only sampling variation is present.) In Cook and Tauchen’s (1984) panel study, for example, spanning eight years and forty-eight states, the standard error of the MLDA effect is about three percentage points. It is only twice as large in the typical quasi-experimental study that analyzed four or five years of data on a single state. Finally, quasi-experimental methods are relatively subjective, with researcher discretion in the state studied, the choice of control group, and the
way in which it is integrated into the analysis (see below).

These technical issues coincide with two longstanding themes in social science that articulate why early studies of the effects of traffic safety legislation could be overly favorable.\(^5\) The first of these, a product of the law and economics literature, builds on Stigler’s “endogeneity of laws” (see Siegelman, 2002, and also Andenaes, 1975). This concept holds that, in order to accurately infer an existing law’s effect on social outcomes, attention must be paid to the circumstances of its adoption. If a law is passed because of a temporary flare-up in an undesirable behavior, if it is associated with other efforts to address the problem in question, if it is adopted as part of a package of broader reforms, traditional statistical methods—regression-based or quasi-experimental—will probably overstate its causal effects. A conversation in a hearing described below nicely illustrates this point—and its subtlety (H4, 1986, p. 13-14):

Rep. Nancy Johnson (R-CT): How do you take into account whether or not, for example, there has been a strong movement within a state among high school students to focus on this problem?

GAO Official: That is the beauty of the control group.

Johnson: In your control group, you have states that have not changed their policy…but do you have all of those same groups that are springing up throughout the Nation, the teenage groups, the MADD parents and so forth?

GAO: [Without directly answering the question, simply notes that some control groups are age-based, such as 21-25 year olds, while others are geographical, such as a neighboring state.]

The second theme, coming out of the criminology and sociology literatures, complements

---

\(^5\) Two other themes explain why traffic safety laws might not be effective, providing a conceptual basis for a failure to reject the standard null hypothesis. One emphasizes the limits of deterrence. In criminology the empirical performance of deterrence theory is mixed; increasing the certainty of punishment—which requires resources—matters much more than increasing its severity (see Lewis, 2009, and many sources therein, and Ross, 1992). The other emphasizes the unanticipated effects of policy, which can serve to offset its effectiveness. Males (1986) and Asch and Levy (1990), for example, argue that higher drinking ages simply postpone drinking onset and the fatalities associated with “inexperienced drinkers.”
and augments the first. It emphasizes the importance of public support in making laws effective, and thus suggests that even if the effects of law in voluntarily-adopting states are accurately estimated, these estimates may not accurately predict the performance of future laws that are not adopted in the same manner.

I have read over most of the papers I have written on this general subject during the past thirty years. In nearly every one of them, I state that the weakest link in attacking this problem has been public support. What we perceive as low-level action against the drunken driver is probably a direct result of lack of public support. We can inform and we can enforce and as a result change behavior through fear for a while. But when we fail to change attitude, regression is bound to occur (Borkenstein, 1985).

[A late-1980s decline in media attention is] one of the reasons I suspect we are seeing a slow-down in the progress, because research continues to show that the most effective laws are those that have a combination of enforcement and repeated publicity (Brian O’Neill, IIHS, H5, Aug., 1988, p. 36.)

[All three law enforcement officers interviewed agreed that] the law is a starting point to make the public more aware of the dangers of drinking and driving, but to be effective the law needed public support. The law enforcement supervisors acknowledged that without public support for the…juvenile drunk driving law, legal sanctions would have a minimal impact in deterring this offense (Lewis, 2009, p. 126-127, discussing zero tolerance laws).

This matters not just in theory, but practice, because many 21 year old MLDAs were adopted because of the Congressional incentive:

[Lack of enforcement] points up some of the concern that I have. I think this is particularly true in the South and the West—there is a resentment of federally imposed standards of that type, and therefore it tends to be discounted at the enforcement level and in the courts. (Jim Burnett, Chairman of the NTSB, H1, 1983, p. 236.)

And sometime, a federal mandate is…more highly resented than any other single factor concerning a law. And in some states the governor’s [highway safety] representative, for example, may not even mention that a requirement is a federal law for fear of raising a red flag. (John Hanna, Deputy Commissioner, Virginia Department of Motor Vehicles, H5, Aug., 1988, p. 48.)

The evidence suggests that these two themes are relevant, and that they help explain the
evolution in findings displayed in Figure 1. As documented by Grant (2011) for three Congressionally-incentivized drunk driving laws and by Miron and Tetelbaum (2009) for the MLDA specifically, fatality changes associated with new laws are consistently larger in those states that adopt them earliest, without being spurred to do so by Congress, even when the estimation method is kept the same.

Miron and Tetelbaum’s results are summarized in Figure 2. The top pane contains estimates of the percentage change in youth traffic fatalities generated by the MLDA of twenty-one, taken from a sequence of individual state time-series regressions, plotted against the year this MLDA was adopted. (These regressions each cover the same span of years and include the same control variables. States maintaining an MLDA of 21 since 1975 are excluded.) The bottom pane contains the cumulative estimate, of all states adopting this MLDA up to that particular point in time. Both panes also contain 95% confidence intervals. The cumulative estimate falls by more than half over the course of the decade spanned in the figure, from 10% to less than 5%. Early-adopting states yield more favorable estimates than late-adopting states do.

Of course, this hindsight was not available in 1984 for the MLDA—the only option was to be judicious in evaluating the evidence:

It seems...clear that establishing a 21-year minimum nationwide drinking age would have a salutary impact on drunk driving statistics, although the extent of the prospective improvement is far from certain. (Sen. Charles Mathias, R-MD, H5, 1984, p. S8226.)

I suspect that you have had some testimony about the overwhelming evidence statistically about this, and I suspect that I am hindered by my 35 years as a researcher when I say that the research is not as overwhelming as we would like it to be. (Morris Chafetz, former director of the NIAAA and member, Presidential Commission on Drunk Driving, H2, 1984, p. 71.)

The results to date of studies increasing (sic) the drinking age have generally been favorable. However, these laws have been in place for only a short time. During that time, other factors which could produce a reduction in accidents have been
present…the question of whether increasing the legal age of purchase will reduce accidents remains to be proven when longer experience with these higher age laws generates sufficient data for a more definitive analysis of impact, from which the effect of transient economic factors can be eliminated. (*Alcohol and Highway Safety: A Review of the State of Knowledge*, 1984, p. 49).6

In the Congressional analysis of the MLDA, which prevailed—this judiciousness, or the optimism of its supporters? We investigate this issue in the next section.

**Section III. The Evaluation of the Evidence in Four Congressional Hearings.**

To examine how the political theatre evaluated the evidence on this issue, we rely on a set of Congressional subcommittee hearings. There a range of witnesses spoke extensively about the evidence on the effects of a raised MLDA and responded to a wide variety of questions. The studies discussed in these hearings are more numerous than those cited (less frequently) on the floors of the House and Senate, while the positions argued by the same individual or organization varied little across time. These hearings thus provide a reasonably detailed and comprehensive record of the various perspectives on the evidence, the way in which these perspectives were presented and examined, and the political and technical skill of the participants.

The first hearing was held in October, 1983, by the House Subcommittee on Commerce, Transportation, and Tourism, one month after a bill was introduced in the House to directly establish an MLDA of 21 nationwide. This well-attended hearing, held when there was increasing momentum for federal action but ambiguity about the form that action might take, featured an exhaustive witness list and active participation from committee members, who

---

6 Despite its relevance, this NHTSA publication in general, and this conclusion specifically, was wholly ignored in all debate and hearings, by NHTSA and everyone else, both before and after the NMDAA became law. It and a later edition are discussed below.
peppered the witnesses with wide array of questions.

The next two hearings were held by the Senate Subcommittees on Surface Transportation and Drug and Alcohol Abuse in June, 1984. These were pro forma, as legislative and executive support for the NMDAA was already established. No members besides the chair attended. Each witness read a prepared statement, but there were few questions.

The last hearing occurred in September, 1986, before the Subcommittee of Investigations and Oversight, House Committee on Public Works and Transportation, to discuss a draft report prepared for its chair, James Oberstar, by the GAO, summarizing and assessing the evidence on the effects of the raised MLDA. The very existence of this hearing testifies to hastyness evaluating the evidence prior to passing the NMDAA:

Congress did take an action in 1984, admittedly without...full committee exploration of the issue, but just on the basis of data at hand, and Congress acted. All right. Now we are trying to come back and analyze the benefits of that action. (Rep. James Oberstar, D-MN, H1, 1986, p. 200.)

No other hearing focused on the MLDA like these four did. However, we did also review several others, held before 1983 or after 1986, that were peripherally related to the issue, along with the relevant Congressional debate, all of which are listed in the chronology in the Appendix.

Despite this variety of purpose, each hearing’s format and participants were similar. Witnesses appeared in groups, read a prepared statement and answered questions. Some questions were probing, to test the accuracy of a claim; others were inquisitive, to gather more information; still others were speculative, to consider a new idea. Witness groups tended to be homogenous: pro-MLDA testimony followed by anti-MLDA testimony, or government officials followed by industry groups followed by student groups, each of which had a material or governmental interest in the hearing’s outcome. Disinterested witnesses were arguably limited to a handful of academics, fewer than one-tenth of those testifying. Thus these hearings, while
invariably collegial, were also inherently adversarial.

Advocates of the higher drinking age included the insurer-funded IIHS, elite safety organizations such as the NSC, grassroots advocates such as MADD and parent-teacher associations, along with NHTSA and the NTSB. These organizations all have extensive experience in the policymaking process; several also have analytical skill. NHTSA, for example, manages the FARS (Fatality Analysis Reporting System) data used in many traffic safety analyses, while the NSC edits the well-regarded *Journal of Safety Research* and the IIHS regularly publishes solid quasi-experimental analyses of traffic safety laws in academic journals.

On the other hand, the opposition, mostly groups representing students and the restaurant and beverage industries, possessed less policymaking experience and little analytical skill.

Thus, in contrast to academia, the evidence on the effects of the MLDA was assessed in an adversarial, political environment under significant time pressure, between two sides matched in their passion about the issue but unequal in technical skill and political experience. For the least disinterested participants, these centripetal forces did not encourage judiciousness, but rather its opposite: selective citation of the evidence and flexible standards as to what constituted evidence.

This selectivity is well-illustrated by focusing on the work of one influential researcher, Alexander Wagenaar. Wagenaar (1981) found that in the year after Michigan raised its drinking age from 18 to 21, in December, 1978, crashes involving 18-20 year old drivers that police reported had been drinking fell by 31%. Because police-reported drinking can be unreliable, a

---

7 The NTSB’s advocacy was unusual. Its specialty is investigating the causes of particular accidents in detail, not assessing the merits of traffic safety legislation. A review of its publications, available online, confirms that the MLDA is the such only such law it has forcefully advocated. While NHTSA consistently argued that raising the MLDA dramatically improved traffic safety, prior to June, 1984, it also argued that the decision to do so should be left to the states, consistent with the views of the President.
common “three-factor surrogate” was also analyzed; it fell by 18%. There was little change in control groups. A later study (Wagenaar, 1983) found that in Maine, which raised its drinking age from 18 to 20 in October, 1977, there was a slight increase in crashes by affected drivers with police-reported drinking, but a 19% reduction in the three-factor surrogate. On the pro-raised-MLDA side, the largest of these four numbers, 31%, was widely cited (it is one of the largest bubbles in Figure 3, below), while the others were never mentioned. The other side, in contrast, also cited one of Wagenaar’s findings—the increase in police-reported drinking crashes after the MLDA rose in Maine.

Coupled to this, also on both sides, were uncontrolled comparisons or anecdotes of dubious inferential value as to the effect of a raised MLDA (H1, 1983, p. 140; H2, 1984, p. 43):

Dr. Arnold Yeager, Physicians for Automotive Safety: My dental office happens to be on a corner, and kids…when they are drinking beer in their car, like to toss the beer bottles…out of the window. In the last 6 months I have picked up far fewer beer cans since the drinking age in New Jersey has been raised.

Robert Snow, Florida Entertainment and Dining Association: The reason we do not agree [with a raised MLDA] is that it has been a failure in the State of Florida, where in 1980 there were 19 fatalities of 18-year-olds when it was legal to drink. In 1982, the last reporting year, there was an increase of 20 percent of those 18-year-old fatalities when it was illegal to drink.

For these participants, all this evidence supported divergent, intransigent assessments, which remained even after the NMDAA was law (H4, 1986, pp. 51, 174, 186):

Allan Williams, VP for Research, IIHS: There is no question that raising the alcohol purchasing age results in fewer alcohol-related motor vehicle crash deaths and injuries in this high-risk group. It reduces them by 10-20%, and it does so year after year…If anything, too much research has been done on this topic.

Michael Birkley, Board Member, National Licensed Beverage Association: Despite the frequently recurring theme in popular accounts of selected studies, we have found no consistently reliable basis for the conclusion that raising the legal drinking age has, can, or is even likely to save lives among the affected age group in any jurisdiction. In our opinion, none of the so-called drinking age impact studies conducted to date are capable of supporting such a conclusion.
These excesses of an adversarial system are to some extent unavoidable, an inevitable consequence of “the fact that the recognition of any condition as a social problem is a political matter.” But while this may explain the absence of judiciousness from the least disinterested participants, countervailing forces—standards of professionalism among policy analysts, and for some, a public mission—temporize such excesses in the others. Their behavior is shaped more by the technical skills and judgement that they possess and that is required of them in order to be credible, what one might call the supply and demand of judiciousness. It is still not enough.

Most fundamentally, the technical skills on which judiciousness is based are not required in order for key witnesses to be credible. NHTSA, the NTSB, and the NSC have natural institutional credibility before Congress, while others, including former government officials and (to a lesser extent) representatives of the IIHS, based in metropolitan Washington, D.C., have personal credibility developed through their involvement with policymakers on a wide range of traffic safety issues. (Innvaer et al., 2002, document that personal connections greatly facilitate the use of research in policy formation.) Technical skills only complement and reinforce these other sources of credibility. Thus, there was no imperative for NHTSA to publish (academically) those few analyses of traffic safety laws that have been produced in-house, including three key MLDA studies from the early 1980s discussed below. In a similar vein, the NTSB’s support of a raised MLDA was not based on a formal review of the evidence, and its widely-repeated calculations of lives saved, used to support its recommendation, were erroneous (Males, 1986). The written record contains no hint that any of this mitigated either organization’s credibility.

Judiciousness is also inhibited by the specialization of advocacy and evaluation in a small number of actors that hark from a uniform intellectual tradition. A longstanding divide within traffic safety separates program evaluators, who deliver rapid estimates of the effects of new
laws using quasi-experimental designs, from social scientists (particularly economists), who conduct large-scale, retrospective regression analyses of laws’ long-run effects. NHTSA and the IIHS operate in the first tradition, not surprisingly, but the foundations for judiciousness—skepticism of quasi-experimental analyses of uncontrolled phenomena, and an appreciation of the social science themes discussed above—stem from the second. This intellectual bifurcation mirrors that in the policy sciences in general, generated by cultural and philosophical differences between the two disciplines, which “seem to be too many…to permit cooperation” (Dror, 1971, Chapter 6, and p. 34), and perpetuated by a substantial temporal divide in their studies of any given issue. As Figure 1 indicates, for the MLDA specifically, the quasi-experimental designs favored by program evaluators predated, by a decade, the panel designs favored by economists (Grant, 2011, shows other literatures are similar). The appearance of the social science themes discussed above takes even longer;\(^8\) by this time program evaluators’ attention has shifted to other topics.

A final contributor is the near-absence of academics, who tend to be more judicious by nature and training (this comes through clearly in the hearings), from testimony. This is partly a matter of specialization, and partly cultural:

> Increasingly, public debates about [modern policy issues] resemble adversary proceedings in a court of law, but with an important difference—the lack of generally accepted rules of evidence. Some participants are able to take advantage of the relative informality of the process, but to scientists even codified adversary procedures seem inappropriate and alien to their tradition. In science the issue is not a witness’s credibility but his specific competence…and this is not reliably established by an adversary debate. (Majone, 1989, p. 4)

In fact, for the MLDA, the divergence goes further: we will soon show that surprisingly few

---

\(^8\) These have only recently made forceful appearances in the empirical literature on drunk driving legislation: by Miron and Tetelbaum (2009) for the MLDA, Freeman (2007) for .08 laws, and Lewis (2009) for zero tolerance laws. These studies come twenty-five, seven, and fourteen years after the appropriate Congressional incentive was enacted.
studies cited in Congressional testimony were conducted by academics or published in academia. In consequence, while both of the intellectual foundations for judiciousness were articulated across these hearings, and though some witnesses’ statements did evince such judiciousness, as quoted previously, such statements were rare and easily overlooked. The testimony of the most prestigious participants, such as the leaders of NHTSA and the NTSB, was unreservedly supportive of the raised MLDA’s effects. Optimism prevailed over judiciousness.

This comes through clearly in the evidence cited in these hearings, shown in the bubble plots in Figure 3. As before, each bubble represents a study, but now its color indicates the authors’ affiliation, while its area is proportional to the number of participants that cited it; studies ultimately published in refereed journals are circumscribed in black. The horizontal axis is the year of release or publication; the vertical axis is the percentage change in fatalities involving affected drivers. Notice, as claimed above, that studies published by academics form a minority of the evidence that is cited; that only about half the evidence cited was ever published; and that none of these published studies were conducted by a government agency.

The top plot in the figure depicts the evidence cited by three high-profile raised-MLDA advocates across the three hearings conducted prior to July, 1984, while the middle plot illustrates the evidence cited by five government agencies or quasi-governmental organizations

9 “Long term improvement in the DWI problem will be achieved only if public attitudes change. Suggestions contrary to this basic proposition serve only to divert attention from workable approaches having the potential for significant benefit to society.... In the field of alcohol control, there have been many examples of programs and control strategies which ultimately have proved ineffective, even when first advocated and employed they seemed to show great promise.” (American Automotive Association, H0, 1982, pp. 671-672)

“Most research published to date is based on faulty premises such as assuming a direct cause and effect relationship between drinking age and crashes without taking into account other variables...[such as] changes in DWI enforcement and increased public education...and covering only short time periods which are inadequate for determining whether changes occurring after a lowering or raising of the drinking age are indicative of long-term effects.” (Ronald Sarasin, Director of Government Relations, National Restaurant Association, H2, 1984, pp. 44-46)
in these same hearings. (The organizations included are identified in the note to the table. Opponents of the raised MLDA, having few supportive studies to cite, mostly referred to the experience of states like Montana and Minnesota, where MLDA increases were not associated with fatality changes.) The weighted or unweighted means or medians of the evidence cited by advocates are all at least 20%. The evidence cited by government is less numerous, less diverse, and even more favorable: these means or medians always exceed 25%. These numbers are similar to those listed by President Reagan in his signing statement, quoted above, but are substantially higher than the best available estimate that could have been obtained at that time.

For this estimate we are indebted to the GAO, which, in 1986, conducted a systematic literature review and evaluation, the subject of the fourth hearing we analyzed. The evidence cited therein, fourteen studies of fatal or injury crashes that met reasonable methodological standards, is listed in the bottom bubble plot in Figure 3. With one small exception, discussed shortly, each of these studies utilizes data that ends in 1982, and so could have been produced before 1984. With three exceptions, identified in the plot, each study was produced by then. The mean and median effect of a raised MLDA across these fourteen studies is 13%.

Most of this evidence comes from single-state studies, but this same figure obtains in the first study to examine evidence from all nineteen states that raised their MLDAs between 1976 and 1982, inclusive: DuMouchel, Williams, and Zador (1987). All other contemporaneous studies that were similarly broad in scope obtained somewhat smaller estimates.10 (These findings suggest that the evidence that was available in 1984 may have come from states where

---

10 Hoskin, Yalung-Matthews, and Carraro’s (1986) quasi-experimental study, Saffer and Grossman’s (1987) pooled TSCS regression analysis, and Hoxie and Skinner’s (1987) traditional panel study obtained estimates of 5%, 8%, and 11%. Each included all states that raised the MLDA over its sample period. DuMouchel, Williams, and Zador’s panel analysis, first available in 1985 and included in the GAO review, spanned the years 1975-1984.
the effect was relatively large. Certainly the distribution of studies across states at that time was not uniform. Of these nineteen states, three—IL, MI, and ME—were studied three times each, while eight others—MD, NJ, RI, GA, OH, TX, CT, and NE—had never been studied."

On the floor of Congress, debate was more wide-ranging and philosophical, and the empirical evidence received only modest attention. Most claims of the NMDAA’s effects devolved to an influential IIHS study of nine MLDA-raising states (Williams et al., 1983, well-represented in Figure 3), which found an average fatality reduction of 28%. These claims were rarely disputed. Citations of this figure collapsed shortly after the NMDAA was passed, however, with the appearance of the studies just mentioned, which each found effects of 13% or less. Twenty years later, as we have mentioned, large-scale panel estimates of the raised MLDA’s long-run effects were half this size. Much of this difference is attributable to the early-adopter effect.

Our documentation of a lack of judiciousness is complete, but not our explanation for it, which is, in a way, reduced form. Two of the three underlying causes, the isolation of policy analysts from social science traditions and from academia, are, for key actors, matters of organizational design. In particular, NHTSA, the federal agency responsible for traffic safety policy, could be structured to possess or develop the human capital required to form a more discerning view of the evidence. In the next section, we show that it has not, and explain why.

IV. NHTSA and the Evaluation of Drunk Driving Countermeasures.

NHTSA, the youngest of the Department of Transportation’s thirteen agencies, was founded in 1970, long after agencies devoted to the safety of aircraft, railroads, and motor
carriers were formed. Its founding coincided with a change in emphasis regarding traffic safety, away from a focus on the provision and safety of roads and the enforcement of basic traffic laws, towards improving the safety of vehicles (Gusfield, 1988), sparked by the 1965 publication of Ralph Nader’s *Unsafe at Any Speed*.

While its mandate has always been to address both vehicle and human factors, the former took precedence in the agency’s early years (Gusfield, 1988). But human factors, particularly drunk driving and restraint use, received much more attention beginning in the 1980s (Zimring, 1988). By 2002, the administrator of NHTSA claimed that human, or behavioral, factors were the predominant cause of traffic fatalities (H6, June, 2002, p. 8), a position still held today (see the March, 2010 hearing discussed below).

Two general approaches, or paradigms, can be adopted toward addressing behavioral factors. One is deterrence-based:

Americans place a high value on individualism. They see the world as malleable to individual will and responsive to choice and moral character. It is to the individual that Americans so frequently look in placing responsibility for social problems. It is the base assumption that supports the great faith we have that punishing the bad guys, the drivers, will deter drinking-driving in a society whose social institutions deter public transportation and support drinking practices with limited constraints (Joseph Gusfield, in Ross, 1992, pp. xi-xii).

The alternative stresses the limits of deterrence and “views drunk driving as a predictable consequence of existing social institutions” (Ross, 1992, p. 167):

My father was an alcoholic. And, boy, I am going to tell you: All I remember from when I was a kid was how alcoholism can just literally destroy a family… But I used to be a police officer years ago, and I guess because of my own background and the experience I had in law enforcement, I am convinced that alcoholism is a sickness that you just cannot cure by tougher penalties. It does not

---

11 The vehicle-factors side of NHTSA differs greatly from the behavioral-factors side. It focuses on regulation and standard setting, absent key technical information that is possessed by the automakers, in the presence of well-matched adversaries, the automobile industry and consumer groups (see Breyer, 1982, and Pecht et al., 2005). The research model discussed herein could have been influenced by the early preeminence of the vehicle side, that is, features suitable to vehicle factors may have been inappropriately applied to behavioral factors as well.
work. It did not work for my dad. And it does not work for anybody else either… So it just seems to me that we ought to be focusing more of our resources on treatment and recovery programs too. (Sen. Ben Nighthorse Campbell, R-CO, H6, Feb., 2002, p. 50.)

The deterrence-based approach prevailed during the 1980s. It was “understandably popular with people who have directly or indirectly, through friends and relatives, experienced harm in the course of alcohol-related crashes”—the natural constituency of drunk driving advocacy groups such as MADD, which rose to prominence during this period (Ross, 1992, p. 176, and multiple sources cited therein). This was buttressed by the concomitant political shift toward conservatism (Reinarman, 1988), the “inevitable change in style that happens when criminal justice initiatives trickle down from elites to the generally conservative crime-control ideology of local America,” and a “hardening of public attitudes about the dangers of driving after drinking…due in part to scientific demonstrations linking elevated blood alcohol with automobile crashes” (Zimring, 1988, pp. 379, 381). In 1984 NHTSA’s Alcohol and Highway Safety (Ch. 6) outlined the key features of their “current approach” to controlling drunk driving. Of the seven points emphasized in that approach, one pertains to seat belt usage, and another to changing societal norms. The other five are deterrence-based (including increasing resources and political pressure for “increased countermeasure activity in the States and communities”).

Ross (1992) deftly analyzes the politics of the deterrence approach, pointing out that it is in many entities’ interest to support it, including that of NHTSA:

Much of the effectiveness of the citizen’s movement [such as MADD] is due to its alliance with the traffic safety establishment. State and federal officials have found the movement useful for demonstrating popular support for statutes and other measures proposed by the safety agencies, while the programs endorsed by the movement have been rendered rational and politically sophisticated in the process. The NHTSA has explicitly recognized the value of this constituency and has taken steps to enlarge and strengthen it (p. 177).

Other groups benefitting from this approach are law enforcement, which gains resources, various
businesses that provide services to drunk driving convicts, and the beverage industry, because this approach emphasizes the culpability of the drinking driver rather than the larger social context which supports the intersection of drinking and driving (also see Reinarman, 1988).

The research model used by NHTSA supports this emphasis on deterrence. Its most essential feature is heavy reliance on contractors. This can easily be documented, because one can search all of NHTSA’s Behavioral Safety Research Reports online. A total of 302 reports fall under the subject heading “Impaired Driving—Alcohol.” NHTSA produced the content of 47 of these reports, with the rest done by contract. Using contractors for much of this research, such as “demonstration projects” of various types, is understandable: a grant is not feasible, as this is not basic research of general interest, nor is in-house production, because of the interdisciplinary nature of the project and its distant location. But using contractors to evaluate traffic safety laws cannot usually be justified this way, because these features do not pertain: this research generally involves analyzing publicly available data with straightforward statistical methods to study a topic of public interest. Yet this work, too, is heavily contractor-based. Of the 25 alcohol-impaired driving research reports that fit the criteria just mentioned, only seven were produced by NHTSA.

On the other hand, a contract allows the sponsoring agency to shape the scope and design of the project and to review the contractor’s final report before deciding whether to release it:

Political actors—advocates and public officials—have [been] moving more toward a “work for hire” model, employing research firms and consultants to give them the research they want, when they want it, on their own terms (Henig, 2008, p. 234).

In this way NHTSA controls methods and (to a lesser extent) the reporting of findings, without developing in-house human capital. As we show shortly, this serves political ends.

Qualitative evidence obtained at the very end of this research project reinforces these
conclusions. We spoke at length about NHTSA with two highly-experienced traffic safety experts: an independent policy advocate and a former high-level administrator in the agency. Both bemoaned the “lack of an independent research capability within the agency,” particularly on the behavioral factors side, and described its origins and consequences as mostly or wholly political, emphasizing that the absence of this research capability benefits certain stakeholders—particularly those, listed above, that gain from policies that are oriented toward deterrence.

In addition, we tried asking NHTSA directly about its use of contracting for behavioral factors research. As a baseline for comparison, we also contacted five other knowledge-producing federal agencies with safety-related orientations (in part or whole): the Consumer Product Safety Commission (CPSC), the Environmental Protection Agency (EPA), the Department of Housing and Urban Development (HUD), the Agency for Healthcare Research and Quality (AHRQ), and the Federal Trade Commission (FTC). In these five agencies, it was not difficult to research a medium-to-high level research administrator, or their representative, who articulated that agency’s research model and give a basic rationale for using that model. In contrast, we were unable to speak or correspond with an appropriate individual within NHTSA, after multiple attempts.12

This emphasis on contracting is complemented by another element of NHTSA’s research model, weak engagement with academia. We searched in the Web of Science to compare

12 The conservations, conducted by phone but for one e-mail exchange, typically lasted half an hour, and were centered around three questions: 1) What general principles are used to determine whether research is done in house, by grant, or by contract?, 2) How does the agency ensure the quality of the research that is produced?, and 3) Do oversight committees evince much concern the topics of the previous two questions? Two attempts were made to reach the administrator in charge of NHTSA’s behavioral factors research; the agency’s communication office was contacted as well. In each case a reply was requested but not received. NHTSA’s National Center for Statistics and Analysis manages the FARS data and produces many reports describing current accident trends, but almost never does the kind of research that is the focus of this paper; neither does the Volpe National Transportation Systems Center.
NHTSA’s academic (article) output from 1995-2010 with that of the other five agencies listed above. The results are listed in the table below:

<table>
<thead>
<tr>
<th>Agency</th>
<th>Articles</th>
<th>Average Number of Authors*</th>
<th>% with Academic Coauthor*</th>
<th>% with Consulting Coauthor*</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHRQ</td>
<td>525</td>
<td>4.1</td>
<td>62</td>
<td>38</td>
</tr>
<tr>
<td>CPSC</td>
<td>59</td>
<td>5.0</td>
<td>28</td>
<td>52</td>
</tr>
<tr>
<td>EPA</td>
<td>91</td>
<td>4.0</td>
<td>60</td>
<td>46</td>
</tr>
<tr>
<td>FTC</td>
<td>195</td>
<td>1.7</td>
<td>30</td>
<td>6</td>
</tr>
<tr>
<td>HUD</td>
<td>71</td>
<td>4.7</td>
<td>60</td>
<td>42</td>
</tr>
<tr>
<td>NHTSA</td>
<td>60</td>
<td>3.7</td>
<td>34</td>
<td>46</td>
</tr>
</tbody>
</table>

* random sample of fifty articles

The FTC and AHRQ publish the most and draw their coauthors predominantly from academia. The CPSC and NHTSA publish the least and draw most of their coauthors from consulting. By these admittedly crude measures, NHTSA’s link to academia is relatively weak.\(^{13}\)

Both features of NHTSA’s research model impede the agency’s exposure to innovation and feedback. This, in turn, facilitates the agency’s emphasis on quasi-experimental studies of short run effects in early-adopting states, both in the studies it sponsors and when assessing the evidence on a law’s effectiveness. As previously noted, these designs tend to generate favorable conclusions, mostly because of the “early-adopter effect” illustrated in Figure 2, but also, to a lesser extent, because the research design allows greater discretion in choosing the treatment and

\(^{13}\) The best example of NHTSA’s academic isolation is the study it relies upon, to this day, to support its claim that an MLDA of 21 reduces fatalities by 13%: a 1985, unpublished, in-house study of the experiences of thirteen states that were early adopters of higher MLDAs (Arnold, 1985—see NHTSA’s March 2005 Traffic Safety Facts Research Note). Ironically, it may no longer physically exist—it is not available in the National Technical Information Service, the Library of Congress, or the Department of Transportation library.
control group and interpreting the results. The weaker findings of later studies are not integrated into NHTSA’s literature evaluations; these studies’ methods, panel regressions of most or all states over long time spans, do not influence the way laws are assessed by NHTSA contractors; the social science themes that preach caution in interpreting the results are neither acknowledged nor heeded.

A second, complementary consequence of this research model is that it limits the amount of in-house human capital available to evaluate the literature or testify before Congress. This is specifically discouraged in the only general government statement we could find on the methods by which the federal government should procure research, a 1962 report to President Kennedy:

> Where management decisions are based substantially on technical judgments…there must be sufficient technical competence within the Government so that outside technical advice does not become technical decisionmaking (sic)…We believe it highly important for the Government to be able to turn to technical advice from its own establishment as well as from outside sources. One major source of this technical knowledge is the Government-operated laboratory or research installation…A strong base of technical knowledge should be continually maintained within the Government service and available for advice to top management (Report to the President on Government Contracting for Research and Development, Bureau of the Budget, May 1962, pp. 9-10).

A reduced supply of such technical knowledge makes NHTSA more susceptible to political influences, which, as we have noted, tend to favor deterrence-based countermeasures. And, without it, the agency never acquires the aforementioned intellectual foundations for judiciousness.

Around the time the NMDAA was passed, NHTSA produced three in-house studies of the effects of raised drinking ages: Maxwell (1981), Klein (1981), and Arnold (1985). Each was included in the GAO review, and each used sound quasi-experimental designs to estimate short run effects in early-adopting states, which ranged from 9-15%. These conclusions were no more or less favorable than in the typical study available at the time, and, with one significant
exception, were well-supported. The evidence produced by NHTSA was not systematically skewed, simply representative of its peers in design and result. Nevertheless, all of these studies are subject to the general weaknesses of the quasi-experimental study designs that were discussed above, and NHTSA’s testimony in the hearings we reviewed indicated, without reservation, that a raised MLDA would lead to large reductions in fatalities.

The NMDAA was not an isolated occurrence. Fifteen years later, NHTSA’s optimism about the effects of deterrence was, if anything, greater, in promoting federal action to encourage states to adopt laws that lower the per se illegal BAC limit to .08 (down, generally, from .10). In 1992 NHTSA recommended states adopt these laws, when there was virtually no evidence on their effectiveness; following President Clinton, it later supported a strong federal incentive operating through the threatened withdrawal of highway funds, just as with the MLDA. This passed Congress in 2000. Though only 3% of all traffic fatalities involve drivers with BACs of .08 or .09, who would be affected by the law, double-digit fatality reductions are found in most of the evidence cited in the NHTSA-produced (1998) “Presidential Initiative for Making .08 BAC the National Legal Limit” advocating this legislation. After identifying a number of methodological problems with several studies of the issue, including some studies sponsored or produced by NHTSA, a 1999 GAO report determined that “the evidence does not conclusively establish that .08 BAC laws, by themselves, result in reductions in the number and severity of

---

14 Klein focuses on fatalities from single-vehicle accidents involving male drivers. He finds that, after Maine’s MLDA was raised from 18 to 20, daytime and nighttime fatalities involving 18-year-olds fell, in about the same proportion, while daytime and nighttime fatalities involving 19-year-olds were both unchanged. Because daytime accidents are treated as a control group, this suggests the law had no effect. But Klein focused instead on a different finding: a 15% reduction in fatalities involving 18-year-old and 19-year-old male drivers in all nighttime accidents, not just those involving a single vehicle; no estimate was obtained for the control group. This generous interpretation was adopted by the GAO and thus included in Figure 3. As Arnold (1985) was not obtainable, as mentioned above, its successor (Womble, 1989) was reviewed.
alcohol-related crashes” and that “NHTSA’s position—that this evidence [on the effectiveness of .08 laws] was conclusive—was overstated.” The only contemporaneous large-scale panel study of the issue, written by an economist (Dee, 2001), was ignored in NHTSA’s later review of the evidence (NHTSA’s “Setting Limits, Saving Lives,” 2000) and in the 2001 edition of *Alcohol and Highway Safety*.15 Dee’s study found that lowering the per se BAC standard from .10 to .08 reduced fatalities by 3%. Subsequent studies of both early-adopting and late-adopting states suggest even that estimate may be too high (for example, Freeman, 2007).

This most recent, 2001 version of *Alcohol and Highway Safety* confirms that NHTSA’s research model, too, is still in effect. It, like some of its predecessors, was produced by (long-time) contractors. Its section on research design (pp. 99-100) discusses quasi-experimental and time-series methods, but not multivariate regression, and virtually no regression-based evaluations of drunk driving legislation are included in its literature reviews. Its conclusion identifies two groups of drunk driving countermeasures: one composed of five items “that have shown promise but for which evaluations of alcohol-crash impact are as yet inconclusive,” the other composed of twelve items “with strong evidence favoring their effectiveness.” No countermeasure was determined to be ineffective. Most of these countermeasures are deterrence-based. In a closing criticism, the contractors note that countermeasures that are not deterrence-based, “focusing on technology, the vehicle, the highway environment, and the more effective control of alcohol consumption…have either been insufficiently developed, insufficiently evaluated, or both” (p. 155). Furthermore, our review of NHTSA’s Behavioral

---

15 While published in 2001, Dee’s paper was available as a manuscript in 2000. The literature review included within his study is mentioned in *Alcohol and Highway Safety*, but its empirical findings are ignored. The only other contemporaneous study by economists, Chaloupka, Saffer, and Grossman (1993), found that .08 laws had no effect on fatalities and was also ignored by NHTSA, the GAO, and *Alcohol and Highway Safety*. 

Factors Research Reports indicates the trend is toward greater reliance on contractors over time. At present, there are no countervailing forces that would materially disturb this equilibrium. Discussions with staff on two oversight committees, representing both parties and both houses of Congress, confirms their lack of concern with such technical matters. Literature reviews—all conducted by program evaluators—have not drawn the technical distinctions between alternative study designs that have been emphasized here. In a recent subcommittee hearing, “Assessing the Effectiveness of the NHTSA’s Highway Traffic Safety Programs,” the only reference to improving the evaluation of behavioral safety initiatives came from the American Automobile Association (Subcommittee on Highways and Transit, House Committee on Transportation and Infrastructure, July 16, 2008, p. 35). A more recent, high-profile hearing (“NHTSA Oversight: The Road Ahead,” Subcommittee on Commerce, Trade, and Consumer Protection, House Committee on Energy and Commerce, March 11, 2010) contained much discussion about the agency’s funding, competence, and openness, but little about human factors and nothing about the evaluation of traffic safety legislation.

V. Conclusions and Prescriptions.

Given the inexorable political nature of traffic safety legislation, what can be done to improve assessments of this legislation at the time they are needed in the political theatre?

One option is to upgrade research methods. “Before/after” quasi-experimental studies, like most panel regressions, do not account for the serial correlation in state fatality rates; time-series analyses, which do, often fail to account for important economic factors (which do not follow simple autoregressive or moving average processes) and are generally applied one state at
a time. The integration of regression-based methods with the use of control groups, while increasingly common among economic traffic safety analyses, are not often used in the early studies that dominate the political debate. Economists’ technical sophistication could be applied to determine the best way to integrate these methods and existing data to form optimum inferences of laws’ effects in those states that adopt them first. One promising approach has been developed by Grant (2010), who shows that analyses of the fraction of fatality-involved drivers that have been drinking can be used to predict the effect of drunk driving laws on fatalities and do not exhibit systematically larger estimates in early-adopting states.

A complementary option is organizational. Enhance the technical research capabilities within NHTSA to incorporate a wider range of methodological prowess, including panel regression methods, and a wider appreciation for the limits of policy that are stressed in significant themes within social science and economics. Fuse greater links between NHTSA’s professional analysts and the academic community, which would include publishing in a range of academic journals and interacting with traffic safety analysts of diverse professional backgrounds. And broaden NHTSA’s research focus to include both the short-run effects of new types of laws and the long-run effects of older laws that have become widely adopted. Research departments at many federal agencies already have this kind of intellectual diversity and these links to academia.

Such changes could easily lead to less enthusiastic empirical and political support for some traffic safety legislation, but that need not mean more traffic fatalities. Every action has an opportunity cost. Political and intellectual capital spent supporting laws that are largely ineffective could be used to seek out and evaluate laws or other, non-deterrence-based mechanisms that may be more effective, and supporting those that ultimately pass the bar.
Table 1. Social Activity Aimed at Reducing Drunk Driving (constructed from Howland, 1988).

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Drunk Driving Groups Founded</th>
<th>Volume of Newspaper Coverage of Drunk Driving</th>
<th>Volume of Periodical Coverage of Drunk Driving</th>
<th>“Legislative Changes to Reduce Drunk Driving”</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978-1981</td>
<td>36</td>
<td>37</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td></td>
<td>17</td>
<td>13</td>
<td>44</td>
</tr>
<tr>
<td>1982</td>
<td>109</td>
<td>81</td>
<td>35</td>
<td>47</td>
</tr>
<tr>
<td>1983</td>
<td>117</td>
<td>169</td>
<td>50</td>
<td>129</td>
</tr>
<tr>
<td>1984</td>
<td>103</td>
<td>162</td>
<td>42</td>
<td>108</td>
</tr>
<tr>
<td>1985</td>
<td>89</td>
<td>76</td>
<td>36</td>
<td>223</td>
</tr>
<tr>
<td>1986</td>
<td></td>
<td>45</td>
<td>9</td>
<td>178</td>
</tr>
</tbody>
</table>

Note: Newspapers include the *New York Times*, the *Washington Post*, the *Los Angeles Times*, and *The Wall Street Journal*. Periodical volume comes from the Magazine Index.
Figure 1. Bubble Plot of Academic Studies of the MLDA (from Grant, 2011).

Note: Black-ringed bubbles are supported by external funding. The year is the year of publication. The estimate of magnitude is the percentage change in the fatality or crash-involvement measure. The volume of the bubble is proportional to the number of citations it received in Google Scholar as of June 2009.
Figure 2. The “Early-Adopter Effect” (based on the findings of Miron and Tetelbaum, 2009). The top graph contains individual state estimates, and the bottom graph cumulative estimates.
Figure 3. Evidence Cited across Four Hearings on the Effect of a Raised MLDA on Crashes.


Note: The advocates are MADD, the American Automobile Association, and the IIHS. The government organizations are NHTSA, the NTSB, the NSC, the National Association of Governor’s Highway Safety Representatives, and the Presidential Commission on Drunk Driving. The area of each bubble is proportional to the number of organizations citing that study in the hearings indicated. Bubble colors indicate the affiliation of the authors of each study; “raw statistics” indicates the organization simply cited the change in crashes or fatalities after the adoption of a law, without referring to any formal study. Bubbles ringed in black indicate studies published in a refereed journal. For such studies, the horizontal axis indicates the publication year; for the others it is the year the study was completed.

REFERENCES


Grant, D. The dynamics of drunk driving in the U.S. Manuscript, Sam Houston State University, 2010.

----. Policy analysis and policy adoption: a study of three national drunk driving initiatives. Manuscript, Sam Houston State University, 2011.


Appendix: A Chronology of the Minimum Drinking-Age Issue
(excerpted almost verbatim from Appendix IV of GAO, 1987, with minor edits and additions; hearings or Congressional debate referenced in the text of this paper are italicized)

Jan. 5, 1933. Ratification of the 21st amendment repealed prohibition and granted the states substantial power to regulate the purchase and possession of liquor within a state.

Sept. 9, 1966. Enactment of the Highway Safety Act of 1966 (Public Law 89-564) provided the first major impetus for federal involvement in drinking and driving by requiring the Department of Transportation (DOT) to establish uniform safety standards for state highway safety programs and to provide funds to carry out such programs.

June 1967. The DOT issued its “Alcohol in Relation to Highway Safety Standard,” to broaden the scope and number of activities directed at reducing alcohol-related accidents.


July 1971. Ratification of the 26th amendment, extending the right to vote to 18-year-olds, helped prompt 29 states to lower their minimum drinking ages in the early 1970’s.

1973. NHTSA agreed by contract with the University of Michigan Highway Safety Research Institute to scientifically analyze the effects of lowering the legal drinking age from 21 to 18 on youths involved in crashes. The report showed a 10%-26% increase in crash involvement between 1968 and 1971.


1976. From this year on, no state lowered its drinking age, partly because of empirical evidence that suggested a link between lowering the drinking age and increased traffic fatalities. Between 1976 and 1980, thirteen states raised their drinking ages by at least one year.

April 14, 1982. The president appointed a 32-member commission to study the national problem of drunk driving.

April 27, 1982. H.R. 6170 was introduced by members of the Congress from New Jersey and Maryland and others to encourage the states to strengthen programs to control drunk driving.

April 29, 1982. [H0] The House Subcommittee on Surface Transportation held hearings on H.R. 6170; the legislation was generally supported by both the beverage and insurance industries.

May 12, 1982. H.R. 6170 was incorporated into H.R. 6211, which became the Surface Transportation Assistance Act of 1982.

July 22, 1982. The National Transportation Safety Board recommended a national minimum drinking age of 21.

Oct. 1, 1982. The Senate unanimously approved its counterpart bill to H.R. 6170, and the bill was sent to the president.

Oct. 25, 1982. Enactment of H.R. 6170 as Public Law 97-364 provided for a two-tier incentive grant program to improve traffic safety. The Congress mandated that the secretary of the Department of Transportation would consider a state minimum drinking age of 21 as one criterion to be met for supplemental grants.

Nov. 30, 1982. House and Senate resolutions were introduced on the legal minimum age for drinking and the purchase of alcohol.

Dec. 13, 1982. The Presidential Commission on Drunk Driving recommended a uniform minimum drinking age of 21 in an interim report intended to allow state legislatures time to consider this recommendation early in their 1983 sessions.

Jan. 6, 1983. The Surface Transportation Assistance Act of 1982 contained a small section strongly encouraging the states to raise the minimum drinking age to 21. On the day the law was enacted, House Concurrent Resolution 23 was introduced, expressing the sense of the Congress that all states should establish a minimum drinking age of 21.

Jan. 27, 1983. A Gallup poll showed that 77 percent of Americans supported a uniform drinking age of 21 for all states.

April 7, 1983. H.R. 2441 was introduced by a member of the Congress from Illinois to prohibit the use of federal highway funds by states whose minimum drinking age was lower than 21.

April 20-21, 1983. Senators from Missouri, Oregon, and Rhode Island introduced S. 1108, the Highway Safety Act of 1983, which provided more incentive grants to states for efforts to deter drunk driving. The bill was never voted out of committee. A member of the Congress from California introduced H.R. 2693, a counterpart bill to S. 1108.

May 6, 1983. A Senator from Pennsylvania introduced Concurrent Resolution 32 to express the sentiment of the Congress that all states should establish a minimum drinking age of 21.

Sept. 13, 1983. Members of the Congress introduced H.R. 3870, a bill to prohibit the sale of alcoholic beverages to persons under 21 years of age under certain conditions.


Nov. 1983. The Presidential Commission on Drunk Driving issued its final report, keeping the recommendation for a uniform minimum drinking age of 21 for the purchase and public possession of all alcoholic beverages.

Jan. 1984. The National Safety Council supported the formation of an organization to follow up on the work of the Presidential Commission, called the National Commission Against Drunk Driving. Also, the president publicly rejected the support of the Presidential Commission on Drunk Driving for a uniform minimum drinking age of 21.

Feb. 7, 1984. Several senators introduced S. 2263, the Uniform Minimum Drinking Age Act, to amend the Surface Transportation Assistance Act of 1982 by reducing the amount of federal highway aid for states that do not enact a legal minimum drinking age of 21.


Feb. and March 1984. The House Subcommittee on Surface Transportation held hearings on surface transportation issues, including a discussion of the drinking-age issue.

April 5, 1984. Members of Congress introduced H.R. 5383, a bill to reduce a state’s apportionment for federal aid for highways in specific fiscal years for states with drinking ages below 21.


May 24, 1984. Senators from New Jersey and Rhode Island introduced S. 2719 as a revision of S. 2263, a counterpart to H.R. 5383, and an attachment to H.R. 4616, the Child Safety Restraint Act.

June 7, 1984. The House approved H.R. 5383 as an amendment to H.R. 5504, which would reduce federal highway funds by 5 percent in fiscal year 1987 and 10 percent in fiscal year 1988 for states not enacting a minimum drinking age of 21.

June 13, 1984. The administration reversed its position on the minimum drinking-age issue through support of H.R. 4616 from the secretary of the Department of Transportation.

June 14, 1984. The Senate Subcommittee on Surface Transportation held hearings on measures to combat drunk driving.

June 19, 1984. The Senate Subcommittee on Alcoholism and Drug Abuse held hearings on a national minimum drinking age.

June 26, 1984. The Senate passed S. 1948 by a vote of 81-16, as an attachment to H.R. 4616. The Senate then passed its version of H.R. 4616 by a voice vote.

June 27, 1984. The House cleared the Senate version of H.R. 4616, including H.R. 5383.

July 6, 1984. The Senate version of H.R. 4616 was approved and sent to the president.

July 17, 1984. The Child Safety Restraint Act (H.R. 4616), which included legislation for a national minimum drinking age of 21, was signed into law (Public Law 98-363) amending the Surface Transportation Assistance Act of 1982. This act was strongly lobbied for by the Mothers Against Drunk Driving, the Parent Teachers Association, the National Safety Council, the National Council on Alcoholism, and the insurance industry.

Sept. 21, 1984. South Dakota brought an action against the secretary of the Department of Transportation in the U.S. District Court for the District of South Dakota, asking the court to declare the uniform national drinking age sanction of the Surface Transportation Assistance Act of 1982 unconstitutional, on the grounds that it violated the 10th and 21st amendments of the U.S. constitution.

May 3, 1985. The U.S. District Court issued a memorandum opinion and judgment dismissing the South Dakota case against the national drinking-age legislation.
May 16, 1985. Members of the Congress from Louisiana and Vermont introduced H.R. 2537 to apportion federal highway funds withheld from states for falling to establish a minimum drinking age of 21 if certain alcohol-related traffic fatalities are significantly reduced.

June 3, 1985. A member of the Congress from Louisiana introduced H.R. 2645 to repeal the national minimum drinking-age law.

June 26, 1985. South Dakota appealed the District Court’s decision to the Court of Appeals for the Eighth Circuit, contending again that the 10th and 21st amendments were violated by the national drinking-age legislation. Nine other non-complying states supported South Dakota’s appeal.

July 11, 1985. Senators from Missouri and New Jersey introduced S. 1428, to make permanent the withholding of 10 percent of the apportionment from the Highway Trust Fund to states that have not adopted the national minimum drinking age.

Sept. 27, 1985. NHTSA and the Federal Highway Administration issued a notice of proposed rulemaking to implement section 6 of Public Law 98-363, which refers to the withholding of federal-aid highway funds.

Oct. 21, 1985. The Chair of the House Subcommittee of Investigations and Oversight, Committee on Public Works and Transportation, asks the GAO to review “existing evaluation[s] of drinking age laws to determine the extent to which they provide empirical support for federal and state initiatives to change the legal drinking age.”

Nov. 12, 1985. S.1428 was amended to S. 1730, the Consolidated Budget Reconciliation Act.

Dec. 20, 1985. S. 1730 was folded into H.R. 3128, the Budget Reconciliation Act, which did not pass but was carried over into the next year.

April 7, 1986. The president signed the Budget Reconciliation Act, which made permanent the withholding of 10 percent of federal highway funds from states not complying with a uniform drinking age.

May 21, 1986. The court of appeals for the eighth circuit affirmed the district court’s dismissal of South Dakota’s complaint challenging the constitutionality of the national drinking-age legislation.

Sept. 16, 1986. Relying on an early draft of the GAO report, the House Subcommittee of Investigations and Oversight, Committee on Public Works and Transportation, conducts hearings to assess evidence on the efficacy of minimum drinking age laws.


Feb. 27, June 27, 2002. The Senate Subcommittee on Transportation and Related Agencies, followed by the House Subcommittee on Highways and Transportation, hold hearings on various traffic safety related issues.