MIT CEEPR

MIT Center for Energy and Environmental Policy Research

Working Paper Series

Are Consumers Poorly-Informed about Fuel Economy? Evidence from Two Experiments

HUNT ALLCOTT AND CHRISTOPHER R. KNITTEL

MARCH 2017

CEEPR WP 2017-008

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Are Consumers Poorly-Informed about Fuel Economy? Evidence from Two Experiments

Hunt Allcott and Christopher Knittel*

January 13, 2017

Abstract

It has long been argued that people are poorly-informed about and inattentive to fuel economy when buying cars, and that this causes us to buy low-fuel economy vehicles despite our own best interest. We test this assertion by running two experiments providing fuel economy information to people shopping for new vehicles. We find zero statistical or economic effect of information on average fuel economy of vehicles purchased. In the context of a simple optimal policy model, the estimates suggest that imperfect information and inattention are not valid as significant justifications for fuel economy standards at current or planned levels.

JEL Codes: D12, D83, L15, L91, Q41, Q48.

Keywords: Behavioral public economics, fuel economy standards, field experiments, information provision.

Consumers constantly choose products under imperfect information. Most goods we buy have many attributes, and it is difficult to pay attention to and learn about all of them. This opens the door to the possibility that we might make mistakes: maybe we should have signed up for a better health insurance plan with a wider network and lower copays, and maybe we wouldn't have bought that coffee if we knew how many calories it has. Indeed, there is significant evidence that consumers

*Allcott: New York University, NBER, and E2e. hunt.allcott@nyu.edu. Knittel: MIT Sloan, NBER, and E2e. knittel@mit.edu. We are grateful to Will Tucker, Jamie Kimmel, and others at ideas42 for research management and to Skand Goel for research assistance. We thank Catherine Wolfram and seminar participants at the 2017 ASSA meetings and the University of California Energy Institute for comments. Funding was provided by the Ford-MIT Alliance, and we are grateful to Ford's Emily Kolinski Morris for her collaboration and support of the experiments. Notwithstanding, Ford had no control over the data, analysis, interpretation, editorial content, or other aspects of this paper. This RCT was registered in the American Economic Association Registry for randomized control trials under trial number AEARCTR-0001421. Screen shots of the interventions and code to replicate the analysis are available from Hunt Allcott's website: https://sites.google.com/site/allcott/research.

can make systematic mistakes when evaluating products, either due to imperfect information about costs and benefits or by failing to pay attention to some attributes.[1]

These issues are particularly important in the context of buying cars. Academics and policymakers have long argued that consumers are poorly informed and cognitively constrained when evaluating fuel economy. Turrentine and Kurani's (2007, page 1213) structured interviews reveal that "when consumers buy a vehicle, they do not have the basic building blocks of knowledge assumed by the model of economically rational decision-making, and they make large errors estimating gasoline costs and savings over time." Many have further argued that these errors *systematically* bias consumers against high-fuel economy vehicles. For example, Kempton and Montgomery (1982, page 826) describe "folk quantification of energy," arguing that "[measurement inaccuracies] are systematically biased in ways that cause less energy conservation than would be expected by economically rational response to price."[2] Such systematic consumer bias against saving energy would exacerbate environmental externalities from energy use.

These assertions of systematic bias have become a core motivation for Corporate Average Fuel Economy (CAFE) standards, which are a cornerstone of energy and environmental regulation in the United States, Japan, Europe, China, and other countries.[3] The U.S. government's Regulatory Impact Analysis (RIA) for CAFE standards finds that they generate a massive win-win: not only do they reduce externalities, but they also save consumers money. Over 2011-2025, the standards are projected to cost $125 billion, reduce externalities (mostly from climate change, local air pollution, and national energy security) by $61 billion, and reduce private costs (mostly from buying gasoline) by $540 billion (NHTSA 2012). Thus, even ignoring externalities, the regulation generates $415 billion in net private benefits, with a private benefit/cost ratio of better than three-to-one. Net private benefits are almost seven times more important than externalities in justifying the regulation. The large net private benefit implies that there must be some large non-externality market failure that is keeping the private market from generating these results in the absence of CAFE.[4]

[1]See, for example, Abaluck and Gruber (2011), Bollinger, Leslie, and Sorensen (2011), Barber, Odean, and Zheng (2005), Grubb (2009), Handel and Kolstad (2015), Hossain and Morgan (2006), Jensen (2010), Kling *et al.* (2012), and others.

[2]It is easy to find other examples of these arguments. For example, Greene *et al.* (2005, page 758) write that "It could well be that the apparent undervaluing of fuel economy is a result of bounded rational behavior. Consumers may not find it worth the effort to fully investigate the costs and benefits of higher fuel economy." Stern and Aronson (1984, page 36) write that "The low economic cost and easy availability of energy made energy users relatively unaware of energy. As a result, energy was not a salient feature in family decisions about purchasing homes and automobiles ... Energy has became invisible to consumers, so that even with some heightened awareness, they may be unable to take effective action." Sanstad and Howarth (1994, page 811) write that "problems of imperfect information and bounded rationality on the part of consumers, for example, may lead real world outcomes to deviate from the dictates of efficient resource allocation."

[3]There is a large literature on various aspects of fuel economy standards in the U.S. – see Austin and Dinan (2005), Goldberg (1998), Jacobsen (2013), and Jacobsen and van Benthem (2015) – and other countries, including Japan (Ito and Sallee 2014), Europe (Reynaert and Sallee 2016), and China (Howell 2016).

[4]"By non-externality," we more precisely mean market failures other than the specific environmental and energy

While some possible market failures are on the supply side – for example, cross-firm spillovers from research and development of fuel economy-improving technologies – significant attention has been focused on demand-side market failures. The U.S. government's RIA argues that information, inattention, "myopia," and other behavioral biases might keep consumers from buying higher-fuel economy vehicles that would save them money in the long run at reasonable discount rates. In particular, the RIA argues that *"consumers might lack the information necessary to estimate the value of future fuel savings,"* and *"when buying vehicles, consumers may focus on visible attributes that convey status, such as size, and pay less attention to attributes such as fuel economy"* (EPA 2012, page 8-7).[5]

This important argument suggests a simple empirical test: does providing fuel economy information cause consumers to buy higher-fuel economy vehicles? If, as the RIAs suggest, consumers are indeed imperfectly informed about fuel costs or do not pay attention to fuel economy in a way that would justify a stringent CAFE standard, then information provision should cause people to buy significantly higher-fuel economy vehicles. If information does not increase the average fuel economy of vehicles purchased, then either some other behavioral bias or market failure must exist, or the RIA must overstate net private benefits. Despite the importance of the CAFE regulation, such an experiment has not previously been carried out, perhaps because of the significant required scale and cost. Any new results to inform the regulation's optimal stringency are particularly timely, as it is up for "midterm review" in 2017.

This paper presents the results of two experiments. The first provided fuel economy information to consumers via in-person intercepts at seven Ford dealerships nationwide. The second provided similar information to consumers in a nationwide online survey panel who reported that they were in the market to buy a new car (of any brand). We later followed up with consumers to record what they bought. Our final samples for the dealership and online experiments comprise 372 and 1,489 vehicle buyers, respectively.

The core of the intervention was to provide individually-tailored annual and lifetime fuel cost information for the several vehicles that the consumer was most closely considering, i.e. his or her "consideration set." To make the cost information more salient, we also provided comparisons to common purchases: "that's the same as it would cost for 182 gallons of milk" or for "8.7 tickets to Hawaii." The interventions did not provide each consumer with all possible information about every

security externalities comprising the $61 billion.

[5]See also the CAFE standard final rule (EPA 2010, page 25510): "In short, the problem is that consumers appear not to purchase products that are in their economic self-interest. There are strong theoretical reasons why this might be so," including that "consumers might lack information" and "the benefits of energy-efficient vehicles may not be sufficiently salient to them at the time of purchase, and the lack of salience might lead consumers to neglect an attribute that it would be in their economic interest to consider."

Gayer (2011) summarizes the arguments: "Energy-efficiency regulations and fuel economy regulations are therefore justified by [cost-benefit analyses] only by presuming that consumers are unable to make market decisions that yield personal savings, that the regulator is able to identify these consumer mistakes, and that the regulator should correct economic harm that people do to themselves."

possible vehicle, nor did they address all possible cognitive biases. They did, however, provide clear information relevant to each consumer's decision, and they did draw attention to fuel economy for at least a few minutes. We designed the interventions to ensure that the treatment group understood and internalized the information provided, and record if they did not. In the dealership experiment, our field staff recorded that about 85 percent of the treatment group completed the intervention. In the online experiment, we ensured completion by requiring all respondents to correctly answer a quiz question before advancing.[6]

In the online experiment, we were able to ask stated preference questions immediately after the intervention. Fuel cost information causes statistically significant but economically small shifts in stated preferences toward higher-fuel economy vehicles in the consideration set, but interestingly, the information robustly causes consumers to *decrease* the general importance they report placing on fuel economy. In the follow-up surveys for both experiments, we find no statistically or economically significant effect of information on average fuel economy of purchased vehicles. With 90 percent confidence, we rule out that the interventions increased fuel economy by more than 1.24 and 0.28 miles per gallon (MPG), respectively, in the dealership and online experiments. There are also statistically zero effects when focusing on subgroups that one might expect to be more influenced by information: those who were less certain about what vehicle they wanted, had spent less time researching, had more variation in fuel economy in their consideration set, or made their purchase sooner after receiving our intervention.

We also lay out a theoretical framework to formalize the potential policy implications of our results. We model vehicle consumers who may misoptimize due to imperfect information and inattention, and a policymaker who addresses the resulting distortion by setting a socially optimal fuel economy standard. Paralleling Diamond (1973) and Allcott and Taubinsky (2015), we first show that the optimal fuel economy standard with tradable credits is set at a level such that the credit price equals the average bias of consumers who are marginal to the policy. We then show that under particular homogeneity or orthogonality assumptions, the optimal fuel economy standard increases average fuel economy by the same amount as an intervention that fully informs and draws attention to fuel economy. Put simply, if full information and attention increases fuel economy by Q miles per gallon but it's not feasible to scale up an information provision program nationwide, then the second-best optimal fuel economy standard to address imperfect information and inattention also increases fuel economy by Q MPG. This theoretical result is useful in that it clarifies how our treatment effects might be used to quantitatively inform optimal policy.

Specifically, our 90 percent confidence intervals rule out that the interventions increased fuel economy by more than 1.24 and 0.28 MPG in the two experiments. Estimates are naturally less

[6]To be clear, the argument made by some policymakers is that consumers are poorly-informed despite the existing attempts at information provision, which include required fuel economy information labels on all new vehicles and the Environmental Protection Agency's website www.fueleconomy.gov. This is why both of our interventions go out of their way to "hit people over the head."

precise when re-weighting the samples to match the nationwide population of new car buyers on observables, but the confidence intervals still rule out increases of more than 3.28 and 0.62 MPG. By contrast, CAFE standards are expected to require increases of 5.7 and 16.2 MPG by 2016 and 2025, respectively, relative to 2005 levels, after accounting for various alternative compliance strategies. Thus, the current CAFE standards are perhaps an order of magnitude larger than might be justified by imperfect information and inattention alone.

There are two major concerns with taking our estimates literally for policy analysis. The first is whether our particular interventions permanently removed all informational and attentional biases, and did nothing else. This concern is why we designed the interventions to provide only hard information and to ensure that people understood and internalized that information, and we took steps to minimize demand effects and non-informational persuasion. Notwithstanding, it is unrealistic that any intervention could fully inform consumers about all aspects of fuel costs for all vehicles that they might buy. Furthermore, we find some evidence that consumers forgot the information we provided, as the immediate effects on stated preference in the online experiment are no longer evident in identical stated preference questions in the follow-up survey.

The second concern is that neither of our two samples is representative of the national population. This is why we implemented two experiments in very different populations, and we take simple steps in the analysis such as re-weighting on observables to match the national population means. Because of these two major concerns, we would certainly not use our exact estimates to determine the optimal CAFE standard. But the fact that the proposed CAFE standards increase fuel economy by an order of magnitude more than can be justified by our confidence intervals significantly moves our priors on whether informational and attentional biases should continue to be used as major justifications for the policy.

The paper's main contribution is to provide the first experimental evidence on the effects of fuel economy information on vehicle choices, and to draw out the potential implications for optimal policy. Our work draws on several literatures. First, it is broadly related to randomized evaluations of information provision in a variety of contexts, including Choi *et al.* (2010) and Duflo and Saez (2003) on financial decisions, Bhargava and Manoli (2015) on takeup of social programs, Jin and Sorensen (2006), Kling *et al.* (2012), and Scanlon *et al.* (2002) on health insurance plans, Bollinger, Leslie, and Sorensen (2011) on calorie labels, Dupas (2011) on HIV risk, Hastings and Weinstein (2008) on school choice, Jensen (2010) on the returns to education, Ferraro and Price (2013) on water use, and many others – see Dranove and Jin (2010) for a review. There are several large-sample randomized experiments measuring the effects of energy cost information for durable goods other than cars, including Allcott and Sweeney (2017), Allcott and Taubinsky (2015), Davis and Metcalf (2016), and Newell and Siikamaki (2014), as well as total household energy use, including Allcott (2011), Dolan and Metcalfe (2013), and Jessoe and Rapson (2015).

Second, one might think of energy costs as a potentially "shrouded" product attribute in the sense of Gabaix and Laibson (2006), and information and inattention as one reason why "shrouding" arises. There is thus a connection to the empirical literatures on other types of potentially shrouded attributes, including out-of-pocket health costs (Abaluck and Gruber 2011), mutual fund fees (Barber, Odean, and Zheng 2005), sales taxes (Chetty, Looney, and Kroft 2009), and shipping and handling fees (Hossain and Morgan 2006). An earlier literature on energy efficiency, including Dubin and McFadden (1984) and Hausman (1979), studied similar issues using the framework of "implied discount rates."

Third, our simple model of optimal taxation to address behavioral biases builds on work by Farhi and Gabaix (2015), Gruber and Koszegi (2004), Lockwood and Taubinsky (2017), Mullainathan, Schwartzstein, and Congdon (2012), and O'Donoghue and Rabin (2006). Energy efficiency policy evaluation has been an active sub-field of this literature, including work by Allcott, Mullainathan, and Taubinsky (2014), Allcott and Taubinsky (2015), Heutel (2015), and Tsvetanov and Segerson (2013).

Finally, we are closely connected to the papers estimating behavioral bias in automobile purchases. There is significant disagreement in this literature. A 2010 literature review found 25 studies, of which 12 found that consumers "undervalue" fuel economy, five found that consumers overvalue fuel economy, and eight found no systematic bias (Greene 2010). The recent literature in economics journals includes Allcott (2013), Allcott and Wozny (2014), Busse, Knittel, and Zettelmeyer (2013), Goldberg (1998), Grigolon, Reynaert, and Verboven (2015), and Sallee, West, and Fan (2016). These recent papers use different identification strategies in different samples, and some conclude that there is no systematic consumer bias, while others find mild bias against higher-fuel economy vehicles.

Sections I-VI present the experimental design, data, baseline beliefs about fuel costs, treatment effects, theoretical model of optimal policy, and conclusion, respectively.

I Experimental Design

Both the dealership and online experiments were managed by ideas42, a behavioral economics think tank and consultancy. While the two interventions differed slightly, they both had the same two key goals. The first was to deliver hard information about fuel costs to the treatment group, without cueing the control group to think about fuel economy. The second was to make sure that people understood the interventions, so that null effects could be interpreted as "information didn't matter" instead of "people didn't understand the information" or "the intervention was delivered poorly."

The two experiments had the same structure. Each began with a baseline survey, then the treatment group received fuel economy information. Some months later, we delivered a follow-up survey asking what vehicle consumers had bought.

I.A Dealership Experiment

We implemented the dealership experiment at seven Ford dealerships across the U.S., in Baltimore, Maryland; Broomfield, Colorado; Chattanooga, Tennessee; Naperville, Illinois (near Chicago); North Hills, California (near Los Angeles); Old Bridge Township, New Jersey (near New York City); and Pittsburgh, Pennsylvania. In each case, Ford's corporate office made initial introductions, then ideas42 met with dealership management and recruited them to participate. We approached nine dealerships in different areas of the country chosen for geographic and cultural diversity, and these were the seven that agreed to participate.[7] This high success rate reduces the likelihood of site selection bias (Allcott 2015). Appendix Figure A1 presents a map of the seven dealership locations.

In each dealership, ideas42 hired between one and three research assistants (RAs) to implement the intervention. Ideas42 recruited the RAs through Craigslist and university career services offices. Of the 14 RAs, ten were male and four were female. The median age was 25, with a range from 19 to 60. Nine of the 14 (64 percent) were Caucasian, and the remainder were Indian, Hispanic, and African-American.

Ideas42 trained the RAs using standardized training materials, which included instructions on what to wear and how to engage with customers. Importantly, the RAs were told that their job was to provide information, not to persuade people to buy higher- (or lower-) fuel economy vehicles. For example, the RA training manual stated that "our explicit goal is not to influence consumers to pursue fuel-efficient vehicles. Rather, we are exploring the ways in which the presentation of information affects ultimate purchasing behavior."

The RAs would approach customers in the dealerships and ask them if they were interested in a gift card in exchange for participating in a "survey."[8] If they refused, the RA would record the refusal. The RAs recorded visually-observable demographic information (gender, approximate age, and race) for all people they approached.

For customers who agreed to participate, the RAs would engage them with an iPad app that asked baseline survey questions, randomized them into treatment and control, and delivered the intervention. The iPad app was designed by a private developer hired by ideas42. The baseline survey asked people the make, model, submodel, and model year of their current car and up to three vehicles they were considering purchasing; we refer to these vehicles as the "consideration set." The iPad also asked additional questions, including two questions measuring how far along they are in the purchase process ("how many hours would you say you've spent so far researching what car to buy?" and "how sure are you about what car you will purchase?") and three questions allowing us to calculate annual and "lifetime" fuel costs ("if you purchase a car, how many years do

[7] We failed to engage one dealership in Massachusetts that was under construction, and our Colorado location was a replacement for another Colorado dealership that declined to participate.

[8] For the first few weeks, we did not offer any incentive, and refusals were higher than we wanted. We then experimented with $10 and $25 Amazon or Target gift cards and found that both amounts reduced refusals by a similar amount, so we used $10 gift cards for the rest of the experiment.

you plan to own it?", "how many miles do you expect that your vehicle will be driven each year?", and "what percent of your miles are City vs. Highway?") The baseline survey concluded by asking for contact information.

The iPad randomly assigned half of participants to treatment vs. control groups. For the control group, the intervention ended after the baseline survey. The treatment group first received several additional questions to cue them to start thinking about fuel economy, including asking what they thought the price of gas will be and how much money it will cost to buy gas for each vehicle in the consideration set. We use these fuel cost beliefs in Section III below, along with similar fuel cost belief questions from the follow-up survey.

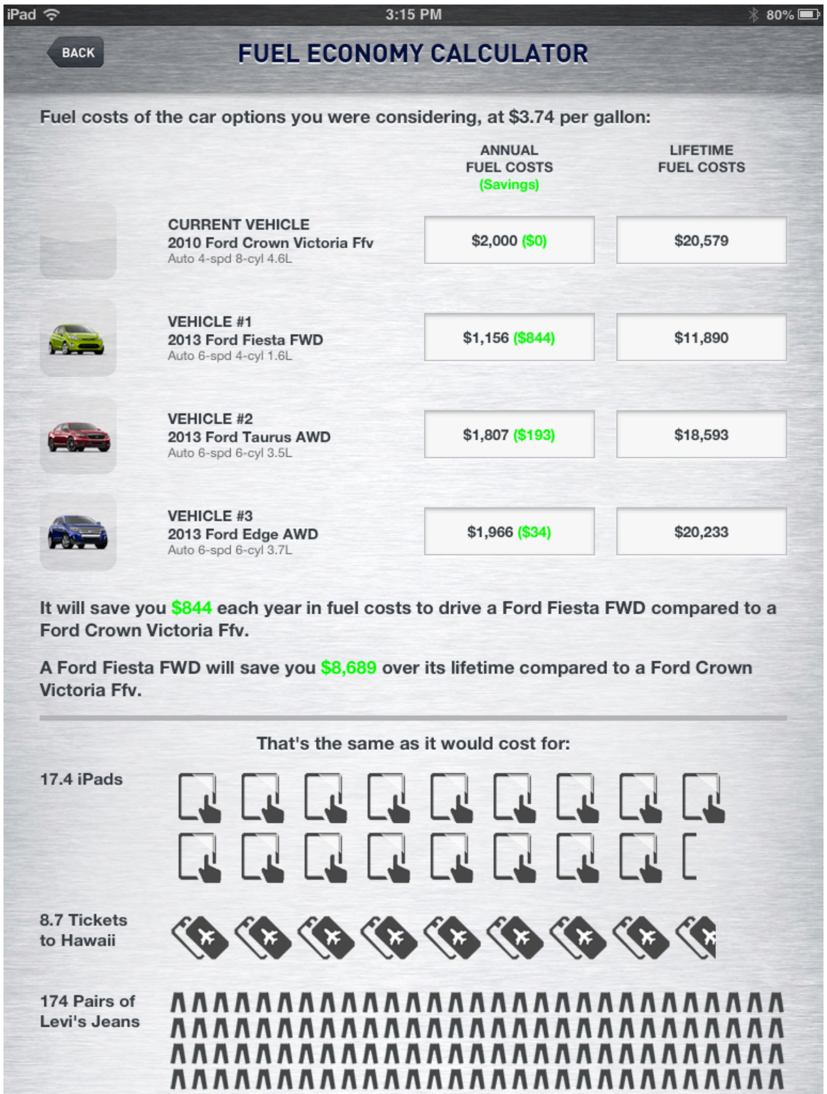
The treatment group then received three informational screens. The first was about MPG Illusion (Larrick and Soll 2008), describing how a two-MPG increase in fuel economy is more valuable when moving from 12 to 14 MPG than when moving from 22 to 24 MPG. The second provided individually-tailored annual and lifetime fuel costs for the consumer's current vehicle and each vehicle in the consideration set, given the participant's self-reported years of ownership, driving patterns, expected gas price. To make these costs salient, the program compared them to other purchases. For example, "A Ford Fiesta will save you $8,689 over its lifetime compared to a Ford Crown Victoria. That's the same as it would cost for 8.7 tickets to Hawaii." Figure 1 presents a picture of this screen. The third screen pointed out that "fuel costs can vary a lot within models," and presented individually-tailored comparisons of annual and lifetime fuel costs for each submodel of each vehicle in the consideration set. After the intervention, we emailed a summary of the information to the participant's email address.

Figure 2 presents a Consort diagram of the dealership experiment and sample sizes. The dealership intercepts happened from December 2012 to April 2014. The follow-up surveys were conducted via phone from August 2013 to September 2014. Of the 3,981 people who were initially approached, 1,740 refused, and 252 accepted but had already purchased a vehicle. Of the remaining 1,989 people, 958 were allocated to treatment and 1,031 to control. Of those allocated to treatment or control, 1,820 people (92 percent) completed the baseline survey.

A subcontractor called QCSS conducted the follow-up survey by phone in three batches: August 2013, January-April 2014, and August-September 2014. There was significant attrition between the baseline and follow-up surveys – some people gave incorrect phone numbers, and many others did not answer the phone. Of those who completed the baseline survey, 398 people (22 percent) completed the follow-up survey. While high, this attrition rate was not unexpected, and 22 percent is a relatively high completion rate for a phone survey. Twenty-six people had not purchased a new vehicle, leaving a final sample of 372 for our treatment effect estimates.

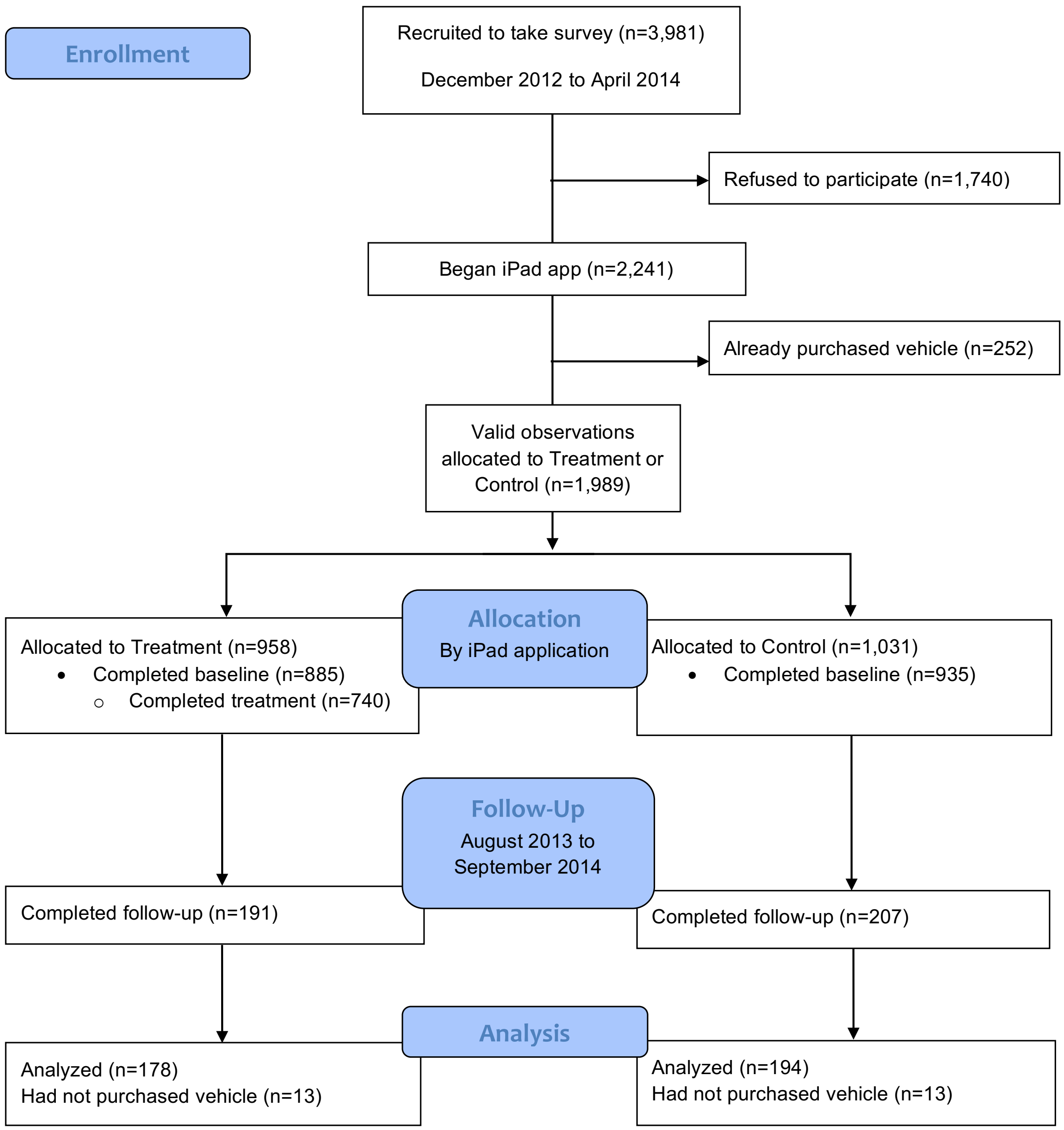
Especially given that we will find a null effect, it is crucial to establish the extent to which the treatment group engaged with and understood the informational intervention. We designed the iPad app to measure completion of the treatment in two ways. First, the participants had to click a "Completed" button at the bottom of the Fuel Economy Calculator screen (the top of which is pictured in Figure 1) in order to advance to the final informational screen. Second, after the intercept was over, the iPad asked the RA, "Did they complete the information intervention?" Of the treatment group consumers who also completed the follow-up survey and thus enter our treatment effect estimates, 87 percent clicked "Completed," and the RAs reported that 85 percent completed the information.

Figure 1: **Dealership Treatment Screen**



Notes: This is a screen capture from part of the iPad-based dealership informational intervention. Vehicles #1, #2, and #3 were those that the participant had said he/she was considering purchasing, and fuel costs were based on self-reported driving patterns and expected gas prices.

Figure 2: **Dealership Experiment Consort Diagram**



RA comments recorded in the iPads suggest that for the 13 to 15 percent of the treatment group that did not complete the intervention, there were two main reasons: distraction (example: "were in a hurry to leave the dealership") and indifference (example: "was not very concerned with fuel efficiency, was looking to purchase a new Mustang for enjoyment"). If non-completion is driven by distraction, we should think of our treatment effects estimates as intent-to-treat, and the local average treatment effect would be 1/0.85 to 1/0.87 times larger. On the other hand, if non-completion is because people are already well-informed or know that their purchases will be unaffected by information, our estimates would reflect average treatment effects.

In the follow-up survey, we also asked, "did you receive information from our researchers about the gasoline costs for different vehicles you were considering?" We would not expect the full treatment group to say "yes," both because they might have forgotten in the months since the dealership interaction, and because someone else in the household could have spoken with the RA. We also might expect some people in the control group to incorrectly recall the interaction. We find that 48 percent of the treatment group recalls receiving information many months later, against 16 percent of the control group.

I.B Online Experiment

For the online experiment, we recruited subjects using the ResearchNow market research panel. The ResearchNow panel includes approximately six million members worldwide, who have been recruited by email, online marketing, and customer loyalty programs. Each panelist provides basic demographics upon enrollment, then takes up to six surveys per year. They receive incentives of approximately $1 to $5 per survey, plus prizes. We began with a sub-sample that were U.S. residents at least 18 years old who reported that they are intending to purchase a car within the next six months.

The online experiment paralleled the dealership experiment, with similar baseline survey, informational interventions, and a later follow-up survey. As in the dealership experiment, we elicited beliefs about annual fuel costs for each vehicle in the consideration set, in both the baseline and follow-up surveys. However, the online experiment offered us a more captive audience and more experimental control, which allowed us to make improvements that were not feasible in the dealership experiment. One improvement was that we could ask additional questions. In the initial survey, before and after the informational interventions, we asked participants the probability that they would buy their first- vs. second-choice vehicles if they had to choose between only those two vehicles, using a slider from 0 to 100 percent. Also immediately after the informational interventions and on the follow-up survey, we asked participants to rate the importance of five attributes on a scale of 1 to 10, as well as how much participants would be willing to pay for four additional features. These questions allow us to construct stated preference measures of the intervention's immediate and long-term effects.

The ResearchNow computers assigned 60 percent of people to treatment and 40 percent to control using a modified least-fill algorithm, which we discuss in more depth below. The base treatment was to provide information similar to the dealership experiment iPad, including annual and "lifetime" (over the expected years of ownership) for the first-choice and second-choice vehicles, as well as for the highest-MPG vehicle in the same class as the first choice. Figure 3 presents a picture of the key information treatment screen. As in the dealership experiment, we compared these fuel costs to other tangible purchases: "that's the same as it would cost for 182 gallons of milk" or for "16 weeks of lunch."

Figure 3: **Online Treatment Screen**

Notes: This is a screen capture from part of the online informational intervention. Choices #1 and #2 were the participant's first-choice and second-choice vehicles, and fuel costs were based on self-reported driving patterns and expected gas prices.

Because we had fully-computerized experimental control instead of delivering the treatment through RAs, we decided to implement four information treatment arms instead of just one. The

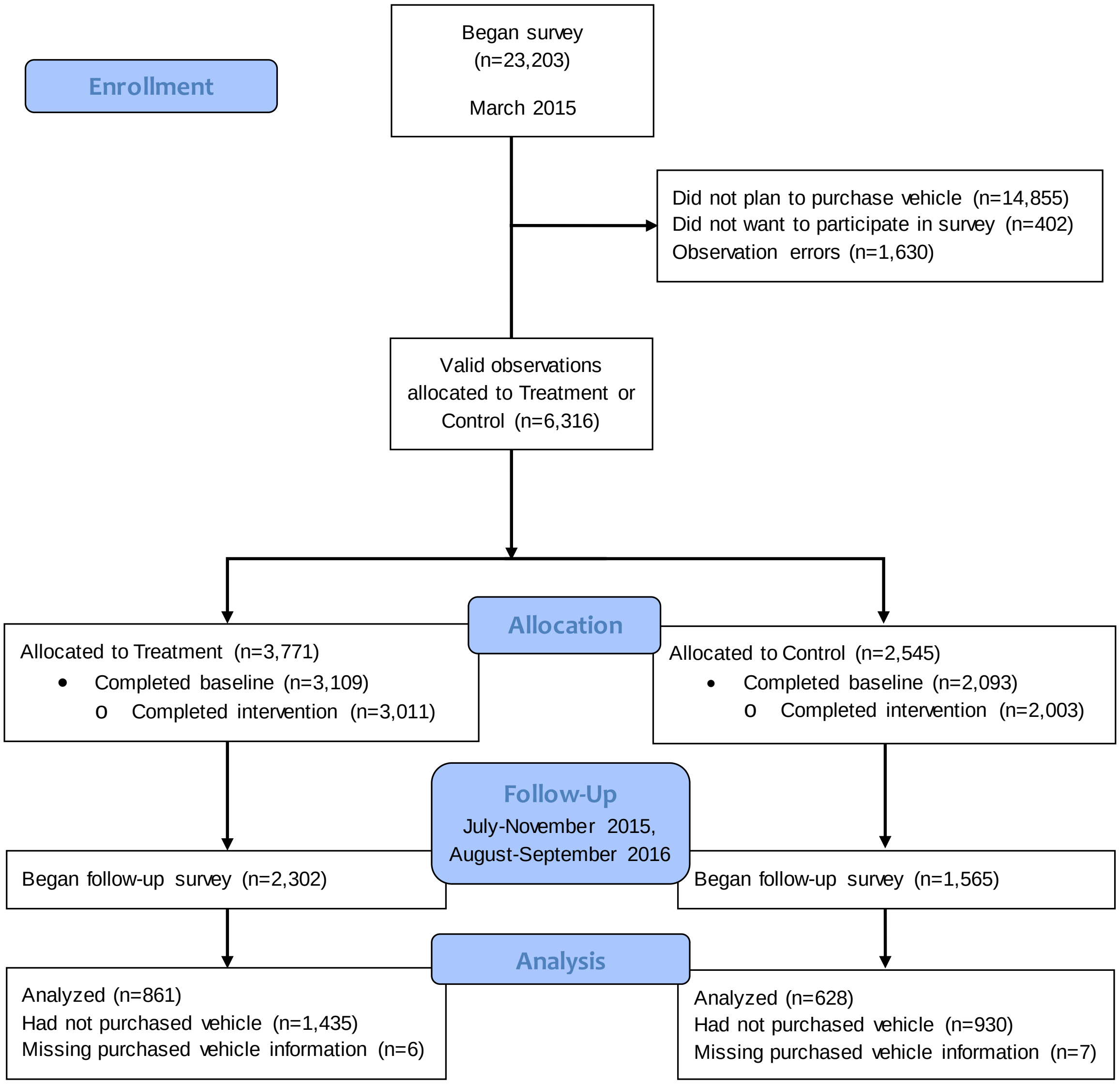
"Base Only" treatment included only the above information, while the other three treatments included additional information. The "Base + Relative" treatment used the self-reported average weekly mileage to compare fuel savings to those that would be obtained at the national average mileage of about 12,000 miles per year. The "Base + Climate" treatment compared the social damages from carbon emissions (monetized at the social cost of carbon) for the same three vehicles as in the Base sub-treatment. The "Full" treatment included all of the Base, Relative, and Climate treatments.

There were also four control groups, each of which paralleled one of the treatment arms in length, graphics, and text, but contained placebo information that was unrelated to fuel economy and would not plausibly affect purchases. One control group was informed about worldwide sales of cars and commercial vehicles in 2007, 2010, and 2013, the second received average vehicle-miles traveled in 2010 vs. 1980, and the third presented the number of cars, trucks, and buses on the road in the U.S. in 1970, 1990, and 2010.

To ensure that people engaged with and understood the information, participants were given a four-part multiple choice question after each of the treatment and control screens. For example, after the base treatment screen in Figure 3, participants were asked, "What is the difference in total fuel costs over [self-reported ownership period] years between the best-in-class MPG model and your first choice vehicle?" Four different answers were presented, only one of which matched the information on the previous screen. 69, 79, and 79 percent of the treatment group answered the Base, Relative, and Environment quiz questions correctly on the first try. 77, 66, and 84 percent of the control group answered the three control group quiz questions correctly on the first try. Every participant was required to answer the questions correctly before advancing.

Figure 4 presents a Consort diagram for the online experiment. The baseline survey and intervention were delivered in March 2015. We conducted the follow-up surveys in two rounds, the first from July to November 2015 and the second in August and September 2016. 6,316 people planned to purchase vehicles and agreed to participate in the survey, of whom 5,014 finished the baseline survey and treatment or control intervention. There is natural attrition over time in the ResearchNow panel, and 3,867 people began the follow-up survey when it was fielded. Of those who began the follow-up survey, 2,378 had not bought a new vehicle or had incomplete data, leaving a final sample of 1,489 people for our treatment effect estimates.

Figure 4: **Online Experiment Consort Diagram**



II Data

II.A Summary Statistics

Table 1 presents summary demographic data for the samples that began the dealership and online experiments – specifically, the samples of valid observations that were randomized into treatment or control. For the dealership experiment, age and the male and Caucasian indicators were coded from the RA notes at the end of the iPad survey, and income is the median income in the consumer's zip code. For the online experiment, demographics are from basic demographics that the respondent provided to ResearchNow upon entering the panel. We impute missing covariates with sample means. See Appendix A for additional details on data preparation.

Given that the dealership sample was recruited at Ford dealerships, it is not surprising that 40 percent of that sample currently drove a Ford, and 67 percent eventually purchased a Ford. By contrast, 12 percent of the online sample currently drove a Ford, and 11 percent purchased a Ford, closely consistent with the national average.

Fuel intensity (in gallons per mile (GPM)) is the inverse of fuel economy (in miles per gallon). For readability, we scale fuel intensity in gallons per 100 miles. The average vehicles use four to five gallons per 100 miles, meaning that they get 20 to 25 miles per gallon. We carry out our full analysis using fuel intensity instead of fuel economy because fuel costs are a key eventual outcome, and fuel costs scale linearly in GPM. "Consideration set fuel intensity" is the mean fuel intensity in the consumer's consideration set.[9]

[9] A small share of vehicles (0.2 to 0.3 percent of purchased and first choice vehicles) are electric. For electric vehicles, the EPA calculates MPG equivalents using the miles a vehicle can travel using the amount of electricity that has the same energy content as a gallon of gasoline. We omit electric vehicles from the descriptive analyses of gasoline cost beliefs, but we include electric vehicles in the treatment effect estimates.

Table 1: **Comparison of Sample Demographics to National Averages**

	(1) Dealership sample	(2) Online sample	(3) National (new car buyers)
Male	0.64	0.60	0.48
	(0.47)	(0.49)	(0.26)
Age	41.47	54.83	54.01
	(12.83)	(13.64)	(13.14)
Caucasian	0.77	0.86	0.91
	(0.41)	(0.35)	(0.29)
Income ($000s)	73.64	121.93	82.08
	(25.71)	(138.33)	(35.68)
Miles driven/year (000s)	15.34	11.68	13.38
	(11.80)	(7.94)	(9.91)
Current vehicle is Ford	0.40	0.12	0.11
	(0.48)	(0.32)	(0.31)
Current fuel intensity (gallons/100 miles)	4.71	4.57	4.58
	(1.15)	(1.08)	(1.50)
Consideration set fuel intensity (gallons/100 miles)	4.34	4.15	0.00
	(1.21)	(0.96)	(0.00)
Purchased fuel intensity (gallons/100 miles)	4.34	4.08	0.00
	(1.28)	(1.00)	(0.00)
Purchased new vehicle	0.67	0.68	1.00
	(0.47)	(0.47)	(0.00)
N	1,989	6,316	18,053

Notes: The first two columns are the final samples used in treatment effects regressions below. "Purchased new vehicle" is an indicator for whether the purchased vehicle is model year 2013 (2015) or later in the dealership (online) sample. The National sample is the sample of households with model year-2008 or later vehicles in the 2009 National Household Travel Survey (NHTS), weighted by the NHTS sample weights. Standard deviations are in parentheses.

The final row reports that 69 to 70 percent of vehicle purchases in the two experiments were "new," as defined by having a model year of 2013 or later (in the dealership experiment) or 2015 or later (in the online experiment). The third column in Table 1 presents the same covariates for the national sample of new car buyers from the 2009 National Household Travel Survey (NHTS), weighted by the NHTS sample weights. For the NHTS, we define "new car buyers" as people who own a model year-2008 or later vehicle in the 2009 survey. Unsurprisingly, neither of our samples is representative of the national population of new car buyers. Interestingly, however, they are selected in opposite ways for some covariates: the online sample is slightly older, significantly wealthier, and drives less than the national comparison group, while the dealership sample is younger, less wealthy, and drives more than the national population. The fact that our samples are selected in different directions on some observables suggests that they might be selected in different directions on unobservables. If this were true, then the treatment effects in our two samples would bound the treatment effect relevant for the national population.

For some regressions, we re-weight the final samples to be nationally-representative on observables using entropy balancing (Hainmueller 2012). We match sample and population means on the six variables in Table 1 that are available in the NHTS: gender, age, the Caucasian indicator, income, miles driven per year, whether the current vehicle is a Ford, and current vehicle fuel intensity. By construction, the mean weight is 1. For the dealership and online samples, respectively, the standard deviations of weights across observations are 1.27 and 0.73, and the maximum observation weights are 11.8 and 9.2.

II.B Balance and Attrition

As mentioned in Section I, ResearchNow allocated observations to the four treatment and four control groups using a modification of the least-fill algorithm.[10] In the standard least-fill algorithm, a survey respondent is allocated to the group with the smallest number of completed surveys. A treatment or control group closes when it reaches the requested sample size, and the survey closes when the last group is full. In this algorithm, between the times when the groups close, group assignment is arbitrary and highly likely to be exogenous, as it depends only on an observation's exact arrival time. Over the full course of the survey, however, group assignment may be less likely to be exogenous, as some treatment or control groups close before others, and different types of people might take the survey earlier vs. later. To address this possible concern, we condition regressions on a set of "treatment group closure time indicators," one for each period between each group closure time.[11] While we include these indicators to ensure that it is most plausible to assume that treatment assignment is unconfounded, it turns out that they have very little impact on the results.

The first eight variables in Table 1 were determined before the information treatment was delivered. Appendix Table A2 shows that F-tests fail to reject that these eight observables are jointly uncorrelated with treatment status. In other words, treatment and control groups are statistically balanced on observables. By chance, however, several individual variables are unbalanced: current vehicle and consideration set fuel intensity in the dealership experiment, and income in the online experiment. We use the eight pre-determined variables as controls to reduce residual variance and ensure conditional exogeneity in treatment effect estimates.

[10]We had instructed ResearchNow to use random assignment, but they did not do this, and we did not discover the discrepancy until we analyzed the data.

[11]We say a "modification" of the least-fill algorithm because there were also some deviations from the above procedure. In particular, had the procedure been followed exactly, the last 20 percent of surveys would all be assigned to a treatment group, as 60 percent of observations were assigned to treatment, versus 40 percent for control. However, ResearchNow modified the algorithm in several ways, and we thus have both treatment and control observations within each of the treatment group closure time indicators.

As we had expected, attrition rates are high. However, this does not appear to threaten internal validity. Appendix Table A3 shows that attrition rates are balanced between treatment and control groups in both experiments, and Appendix Table A4 shows that attrition rates in treatment and control do not differ on observables. On the basis of these results, we proceed with the assumption that treatment assignment is unconfounded.

III Beliefs

Before estimating treatment effects of information, we ask a more basic question: what are consumers' initial beliefs about fuel costs? To do this, we follow Allcott (2013) in constructing valuation ratios, which reflect the ratio of perceived to true fuel cost differences between a pair of vehicles. We define $\tilde{G}_{ij}$ as consumer $i's$ belief about annual gas costs of vehicle j, and G^*_{ij} as the "true" value given the vehicle's fuel economy rating and the consumer's self-reported miles driven, city vs. highway share, and per-gallon gasoline price. For a given vehicle j, consumer i's valuation ratio is

$$\phi_{ij} = \frac{\tilde{G}_{ij}}{G^*_{ij}}. \tag{1}$$

For any pair of vehicles $j \in \{1, 2\}$, consumer i's valuation ratio is

$$\phi_i = \frac{\tilde{G}_{i1} - \tilde{G}_{i2}}{G^*_{i1} - G^*_{i2}}. \tag{2}$$

The valuation ratio measures the share of true fuel costs (or fuel cost differences) that the consumer perceives. The correct benchmark is $\phi = 1$. $\phi > 1$ if the consumer perceives large fuel costs, and $\phi < 1$ if the consumer perceives small fuel costs. Larger $|\phi - 1|$ reflects more "noise" in beliefs.

For example, consider two vehicles, one that gets 25 MPG (4 gallons per 100 miles) and another that gets 20 MPG (5 gallons per 100 miles). For a consumer who expects to drive 10,000 miles per year with a gas price of \$3 per gallon, the two cars would have "true" annual fuel costs $G^*_{i1} = \$1200$ and $G^*_{i2} = \$1500$. If on the survey, the consumer reports $\tilde{G}_{i1} = \$1400$ and $\tilde{G}_{i2} = \$1250$, we would calculate $\phi_i = \frac{1400-1250}{1500-1200} = 0.5$. In other words, the consumer responds as if she recognizes only half of the fuel cost differences between the two vehicles.

The fuel cost beliefs elicited in the surveys are a combination of consumers' actual beliefs plus some survey measurement error. Survey measurement error is especially important due to rounding (most responses are round numbers) and because we did not incentivize correct answers. Appendix Table A5, however, shows that elicited beliefs appear to be meaningful, i.e. not just survey measurement error: the results suggest that ϕ_{ij}, ϕ_i, and $|\phi_i - 1|$ are correlated within individual between the baseline and follow-up surveys, and that people who perceive larger fuel cost differences (lower ϕ_i) also buy higher-MPG vehicles, although the results from the dealership experiment are imprecise due to the smaller sample.

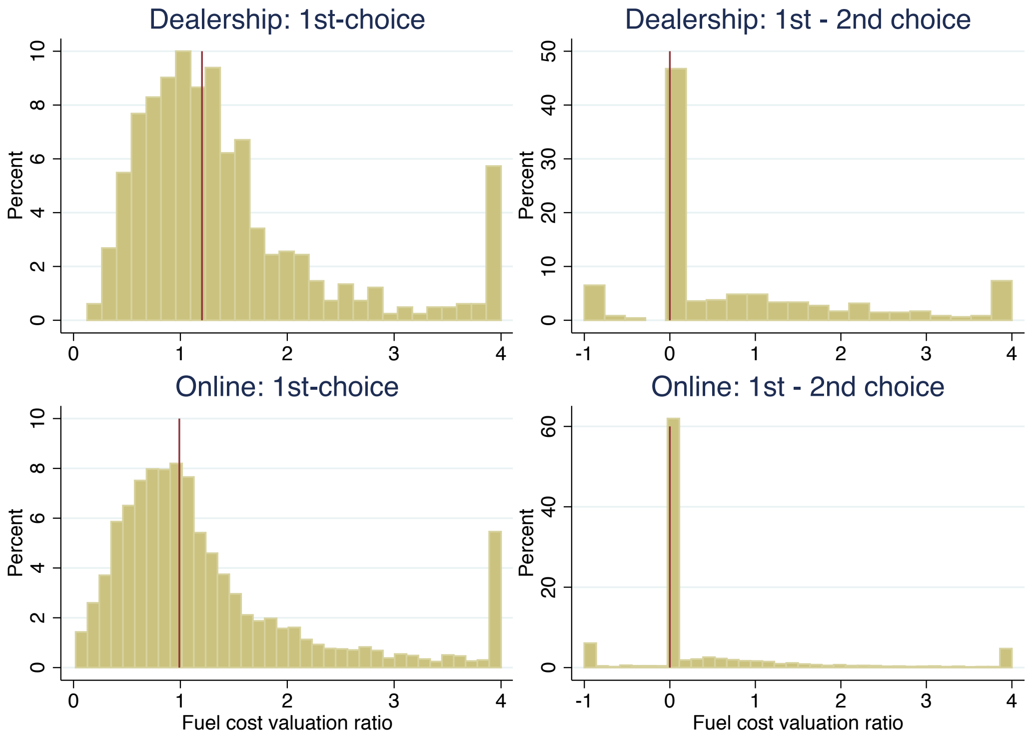
Figure 5 presents the distributions of valuation ratios in the baseline dealership and online surveys. The left panels show ϕ_{ij} from Equation (1) for the first-choice vehicles, while the right panels show ϕ_i from Equation (2) for the first- vs. second-choice vehicles. Since there can be significant variation in ϕ_i, especially for two vehicles with similar fuel economy, we winsorize to the range $-1 \leq \phi \leq 4$.

The figure demonstrates three key results. First, people's reported beliefs are very noisy. Perfectly-reported beliefs would have a point mass at $\phi = 1$. In the dealership and online experiments, respectively, 24 and 32 percent of ϕ_{ij} in the left panels are off by a factor of two or more, i.e. $\phi_{ij} \leq 0.5$ or $\phi_{ij} \geq 2$. This reflects some combination of truly noisy beliefs and survey reporting error.

Second, many people do not correctly report whether their first or second choice vehicle has higher fuel economy, let alone the dollar value of the difference in fuel costs. 45 and 59 percent of respondents in the dealership and online data, respectively, have $\phi_i = 0$, meaning that they reported the same expected fuel costs for vehicles with different fuel economy ratings. In both surveys, eight percent have $\phi_i < 0$, meaning that they have the MPG rankings reversed. Thus, in the dealership and online surveys, respectively, only 47 and 33 percent of people correctly report which of their first- vs. second-choice vehicle has higher fuel economy. This result also reflects some combination of incorrect beliefs and survey reporting error.

Third, it is difficult to argue conclusively whether people systematically overstate or understate fuel costs. The thin vertical lines in Figure 5 mark the median of each distribution. The top left figure shows that the median person in the dealership survey overestimated fuel costs by 20 percent ($\phi_{ij} = 1.2$), which amounts to approximately \$200 per year. The median person in the online survey, by contrast, has $\phi_{ij} = 0.99$. The median $\phi_i = 0$ in both surveys, reflecting the results of the previous paragraph. All four histograms show significant dispersion, making the means harder to interpret.

Figure 5: **Distributions of Fuel Cost Beliefs: Valuation Ratios**



Notes: These figures present the distribution of valuation ratios in the baseline surveys for the dealership and online experiments. The left panels present the valuation ratio from Equation (1) for the first choice vehicles. The right panels present the valuation ratios from Equation (2) for the first- vs. second-choice vehicles. In the right panels, a valuation ratio of zero means that the consumer reported the same expected fuel costs for both vehicles.

IV Empirical Results

We estimate the effects of information by regressing the purchased vehicle fuel intensity on a treatment indicator, controlling for observables. Define Y_i as the fuel intensity of the vehicle purchased by consumer i, measured in gallons per 100 miles. Define T_i as a treatment indicator, and define $\mathbf{X}_i$ as a vector of controls for the eight pre-determined variables in Table 1: gender, age, a Caucasian indicator, natural log of income, miles driven per year, an indicator for whether the current vehicle is a Ford, current vehicle fuel intensity, and consideration set average fuel intensity. The latter two variables soak up a considerable amount of residual variance in Y_i. For the online experiment, $\boldsymbol{X}_i$ also includes the treatment group closure time indicators. The primary estimating equation is

$$Y_i = \tau T_i + \beta \mathbf{X}_i + \varepsilon_i. \tag{3}$$

We first study effects on stated preference questions in the online experiment, both immediately after the intervention and in the follow-up survey. The immediate stated preference questions are useful because they show whether the intervention had any initial impact. By comparing effects on the exact same questions asked months later during the follow-up, we can measure whether the intervention is forgotten. We then estimate effects on the fuel economy of purchased vehicles, for the full sample and then for subgroups that might be more heavily affected.

IV.A Effects on Stated Preference in the Online Experiment

We first show immediate effects on stated preference questions asked just after the online intervention. To increase power, we use the full sample available from the baseline survey, which includes many participants who do not appear in the follow-up survey. Table 2 reports results for three sets of questions. Panel (a) reports estimates of Equation (3) where the dependent variable is the response to the question, "How important to you are each of the following features? (Please rate from 1-10, with 10 being "most important.)" Panel (b) reports estimates where the dependent variable is the answer to the question, "Imagine we could take your most likely choice, the [first-choice vehicle], and change it in particular ways, keeping everything else about the vehicle the same. How much additional money would you be willing to pay for the following?" In both panels, the feature is listed in the column header. Panel (c) presents the expected fuel intensity, i.e. weighted average of the first- and second-choice vehicles, weighted by the post-intervention reported purchase probability. In panel (c), the R^2 is very high, and the estimates are very precise. This is because $\boldsymbol{X}_i$ includes the consideration set average fuel intensity, which is the same as the dependent variable except that it is not weighted by post-intervention reported purchase probability.

Results in Panels (a) and (b) show that the information treatment actually *reduced* the general stated importance of fuel economy. The treatment group rated fuel economy 0.56 points less important on a scale of 1-10 and was willing to pay \$92.19 and \$237.98 less for five and 15 MPG fuel economy improvements, respectively. The treatment also reduced the stated importance of price, although the effect size is less than half of the effect on fuel economy. Preferences for power, leather interior, and sunroof are useful placebo tests, as the intervention did not discuss these issues. As expected, there are no effects on preferences for these attributes.

Table 2: **Immediate Effect of Information on Stated Preference in Online Experiment**

	(1) Power	(2) Fuel economy	(3) Price	(4) Leather interior	(5) Sunroof
Treatment	-0.04	-0.56***	-0.24***	-0.06	0.10
	(0.06)	(0.06)	(0.05)	(0.09)	(0.08)
N	5,036	5,036	5,036	5,036	5,036
R^2	0.04	0.13	0.06	0.07	0.04
Dependent variable mean	6.62	7.68	8.31	4.65	3.80

(a) **Importance of Features, from 1 (Least Important) to 10 (Most Important)**

	(1) Leather interior	(2) 5 MPG improvement	(3) 15 MPG improvement	(4) Power: 0-60 MPH 1 second faster
Treatment	4.49	-92.18***	-237.95***	16.85
	(16.77)	(15.81)	(35.14)	(19.35)
N	4,609	4,512	4,512	4,609
R^2	0.06	0.06	0.07	0.05
Dependent variable mean	380	409	1043	242

(b) **Willingness-to-Pay for Additional Features**

	(1) Expected fuel intensity (gallons/100 miles)
Treatment	-0.032***
	(0.004)
N	5,018
R^2	0.97
Dependent variable mean	4.12

(c) **Expected Fuel Intensity**

Notes: This table presents estimates of Equation (3). The dependent variables in Panel (a) are responses to the question, "How important to you are each of the following features? (Please rate from 1-10, with 10 being "most important.)" Dependent variables in Panel (b) are responses to the question, "Imagine we could take your most likely choice, the [first choice vehicle], and change it in particular ways, keeping everything else about the vehicle the same. How much additional money would you be willing to pay for the following?" In both panels, the feature is listed in the column header. In Panel (c), the dependent variable is the weighted average fuel intensity (in gallons per 100 miles) of the two vehicles in the consideration set, weighted by post-intervention stated purchase probability. Data are from the online experiment, immediately after the treatment and control interventions. All columns control for gender, age, a Caucasian indicator, natural log of income, miles driven per year, an indicator for whether the current vehicle is a Ford, current vehicle fuel intensity, consideration set average fuel intensity, and treatment group closure time indicators. Robust standard errors are in parentheses. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Why might the intervention have reduced the importance of fuel economy? One potential explanation is that people initially overestimated fuel costs and fuel cost differences, and the quantitative information in the treatment helps to correct these biased beliefs. As we saw in Figure 5, however, there is no clear evidence that this is the case for the online experiment sample. Furthermore, we can calculate the actual annual savings from five and 15 MPG fuel economy improvements given each consumer's expected gasoline costs and driving patterns and the MPG rating of the first-choice vehicle. The control group has average willingness-to-pay of \$464 and \$1186 for five and 15 MPG improvements, respectively. The actual *annual* savings are \$266 and \$583. This implies that the control group requires a remarkably fast payback period – approximately two years or less – for fuel economy improvements. It therefore seems unlikely that the control group overestimated the value of fuel economy improvements. Notwithstanding, the results in Panels (a) and (b) are very robust – for example, they are not driven by outliers, and they don't depend on whether or not we include the control variables $\boldsymbol{X}_i$.

Panel (c) of Table 2 shows that the treatment shifted purchase probabilities toward the higher-MPG vehicle in consumers' consideration set. This effect is small: a 25-MPG car has a fuel intensity of 4 gallons per 100 miles, so a decrease of 0.032 represents only a 0.8 percent decrease. In units of fuel economy, this implies moving from 25 to 25.2 miles per gallon.

It need not be surprising that the intervention shifted stated preference toward higher-MPG vehicles in the consideration set while also reducing the stated general importance of fuel economy. As we saw in Figure 5, about two-thirds of online survey respondents do not correctly report which vehicle in their consideration set has higher MPG. Thus, even if the treatment makes fuel economy less important in general, it is still a positive attribute, and the treatment can shift preferences toward higher-MPG vehicles by clarifying which vehicles are in fact higher-MPG. Furthermore, even consumers who do correctly report which vehicle in their consideration has lower fuel costs may be uncertain, the treatment helps make them more certain.

We also asked the same stated preference questions from Panels (a) and (b) on the follow-up survey, which respondents took four to 18 months later. Table 3 parallels Panels (a) and (b) of Table 2, but using these follow-up responses. Only one of the nine variables (importance of price from 1-10) has an effect that is statistically significant with 90 percent confidence. For the fuel economy variables, there are zero remaining statistical effects, and we can reject effects of the sizes reported in Table 2. This suggests that the effects of information wear off over time, perhaps as people forget.

Because these are unincentivized stated preference questions, we are careful to not interpret them too seriously. These results are important, however, because they clearly show that the treatments did have at least some initial impact. Below, we continue by looking at effects on actual vehicle purchases.

Table 3: **Effect of Information on Stated Preference in Online Experiment Follow-Up Survey**

	(1) Power	(2) Fuel economy	(3) Price	(4) Leather interior	(5) Sunroof
Treatment	0.12	-0.10	-0.17*	0.15	0.07
	(0.12)	(0.11)	(0.10)	(0.17)	(0.16)
N	1,542	1,544	1,543	1,542	1,541
R^2	0.03	0.07	0.03	0.05	0.03
Dependent variable mean	6.90	7.76	8.49	4.95	4.02

(a) **Importance of Features, from 1 (Least Important) to 10 (Most Important)**

	(1) Leather Interior	(2) 5 MPG improvement	(3) 15 MPG improvement	(4) Power: 0-60 MPH 1 second faster
Treatment	-37.39	2.67	20.38	13.46
	(29.38)	(23.98)	(56.25)	(27.75)
N	1,359	1,329	1,329	1,359
R^2	0.06	0.04	0.04	0.03
Dependent variable mean	316	346	940	168

(b) **Willingness-to-Pay for Additional Features**

Notes: This table presents estimates of Equation (3). The dependent variables in Panel (a) are responses to the question, "How important to you are each of the following features? (Please rate from 1-10, with 10 being "most important.)" Dependent variables in Panel (b) are responses to the question, "Imagine we could take your most likely choice, the [first choice vehicle], and change it in particular ways, keeping everything else about the vehicle the same. How much additional money would you be willing to pay for the following?" In both panels, the feature is listed in the column header. Data are from the follow-up survey for the online experiment. All columns control for gender, age, a Caucasian indicator, natural log of income, miles driven per year, an indicator for whether the current vehicle is a Ford, current vehicle fuel intensity, consideration set average fuel intensity, and treatment group closure time indicators. Robust standard errors are in parentheses. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

IV.B Effects on Vehicle Purchases

Did the interventions affect only stated preference, or did they also affect actual purchases? Table 4 presents treatment effects on the fuel intensity of purchased vehicles. Columns 1-3 present dealership experiment results, while columns 4-6 present online experiment results. Columns 1 and 4 omit the $\boldsymbol{X}_i$ variables, presenting a simple difference in means. Columns 2 and 5 add $\boldsymbol{X}_i$; the point estimates change little. Columns 3 and 6 are weighted to match U.S. population means, as described in Section II. In all cases, information provision does not statistically significantly affect the average fuel intensity of the vehicles consumers buy.

Are the estimates precise enough to rule out economically significant effects? The bottom row of

Table 4 presents the lower bound of the 90 percent confidence interval of the treatment effect. With equally-weighted observations in columns 2 and 5, the confidence intervals rule out fuel intensity decreases of 0.09 and 0.04 gallons per hundred miles in the dealership and online experiments, respectively. When re-weighted to match the national population, the confidence intervals rule out decreases of 0.50 and 0.08 gallons per hundred miles, respectively. For comparison, for a 25-MPG car, a decrease of 0.1 gallons per 100 miles represents a decrease from 4 to 3.9 gallons per 100 miles, i.e. an increase from 25 to 25.64 miles per gallon.

Table 4: **Effects of Information on Fuel Intensity of Purchased Vehicles**

	(1)	(2)	(3)	(4)	(5)	(6)
		Dealership			Online	
Treatment	0.07	0.11	-0.21	0.06	0.03	0.02
	(0.13)	(0.11)	(0.17)	(0.05)	(0.04)	(0.06)
N	374	374	374	1,489	1,489	1,489
R^2	0.00	0.40	0.30	0.00	0.39	0.38
Dependent variable mean	4.33	4.33	4.33	4.09	4.09	4.09
Controls	No	Yes	Yes	No	Yes	Yes
Weighted	No	No	Yes	No	No	Yes
90% confidence interval lower bound	-0.15	-0.06	-0.48	-0.03	-0.04	-0.08

Notes: This table presents estimates of Equation (3). The dependent variable is the fuel intensity (in gallons per 100 miles) of the vehicle purchased. All columns control for gender, age, a Caucasian indicator, natural log of income, miles driven per year, an indicator for whether the current vehicle is a Ford, current vehicle fuel intensity, and consideration set average fuel intensity. Columns 4-6 also control for treatment group closure time indicators. Samples in columns 3 and 6 are weighted to match the national population of new car buyers. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

As discussed in Section I, the online intervention actually had four separate sub-treatments. Appendix Table A6 presents estimates of Equation (3) for stated preference fuel intensity immediately after the intervention, paralleling Panel (c) of Table 2, and for fuel intensity of purchased vehicles, paralleling column 5 of Table 4. For both outcomes, Wald tests fail to reject that the coefficients on the four sub-treatments are jointly equal. Interestingly, the "Base + Climate" treatment, which included information about both fuel costs and climate change damages, has a statistically positive treatment effect on purchased vehicle fuel intensity, meaning that it caused people to buy statistically lower-fuel economy vehicles. It would be useful to test whether this replicates in other samples.

IV.C Effects in Subgroups

Several hypotheses predict specific subgroups where the treatment effects might be larger or smaller. First, information might have smaller effects on people who are considering vehicles only in a narrow fuel economy range: fuel economy information will likely have smaller effects for a consumer deciding between 22- and 23-MPG vehicles compared to a consumer deciding between a Hummer and a Prius. Second, as suggested by comparing the stated preference results between baseline and follow-up in Tables 2 and 3, the treatment's possible impact may have worn off as people forgot the information. Consumers who bought their new cars sooner after the intervention are less likely to have forgotten. Third, information might be more powerful for people who have done less research and are less sure about what car they want to buy.

Table 5 presents estimates in specific subgroups that, per these hypotheses, might be more responsive. Column 1 re-produces the treatment effect estimate for the full sample. Column 2 considers only consumers with above-median variance of fuel intensity in their consideration set. Column 3 considers only the consumers with below-median time between the intervention and the date of vehicle purchase reported in the follow-up survey. Column 4 drops the approximately half of consumers who report being "almost certain" what vehicle they will purchase, using only consumers who are "fairly sure," "not so sure," or "not at all sure." Column 5 considers only consumers who report having spent less than median time researching what vehicle to buy. In all subgroups, the effects are statistically zero.

Table 5: **Treatment Effects for Subgroups Hypothesized to Be More Responsive**

	(1) Full sample	(2) ≥ Median consideration set MPG variance	(3) ≤ Median time until purchase	(4) Less sure	(5) ≤ Median research time
Treatment	0.11	0.05	0.08	0.22	0.25
	(0.11)	(0.17)	(0.14)	(0.15)	(0.16)
N	374	187	168	185	166
R^2	0.40	0.31	0.47	0.40	0.43
Dependent variable mean	4.33	4.16	4.25	4.24	4.23

(a) **Dealership Experiment**

	(1) Full sample	(2) ≥ Median consideration set MPG variance	(3) ≤ Median time until purchase	(4) Less sure	(5) ≤ Median research time
Treatment	0.03	-0.03	0.01	0.03	0.00
	(0.04)	(0.06)	(0.06)	(0.05)	(0.06)
N	1,489	744	745	1,095	743
R^2	0.39	0.36	0.44	0.35	0.42
Dependent variable mean	4.09	3.93	4.06	4.10	4.07

(b) **Online Experiment**

Notes: This table presents estimates of Equation (3), with samples limited to the subgroups indicated in the column headers. The dependent variable is the fuel intensity (in gallons per 100 miles) of the vehicle purchased. All columns control for gender, age, a Caucasian indicator, natural log of income, miles driven per year, an indicator for whether the current vehicle is a Ford, current vehicle fuel intensity, and consideration set average fuel intensity. Panel (b) also includes treatment group closure time indicators. Robust standard errors are in parentheses. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

V Theoretical Model: Implications for Optimal Policy

V.A Model Setup

We now present a theoretical framework that – under additional structure and assumptions – suggests a potential policy implication of our empirical results. We formalize a model of vehicle buyers who may misoptimize due to imperfect information and inattention, and a social planner who sets an optimal fuel economy standard. This model formalizes the arguments presented in the introduction that imperfect information and inattention cause systematic misoptimization, and that CAFE standards can help address these distortions. The model generalizes the main results of Allcott and Taubinsky (2015) from two choices to $J \geq 2$ choices, which is important for the auto market setting, and clarifies how to analyze fuel economy standards in this framework.

The social planner wants to set the socially-optimal fuel economy standard. Consistent with the current policy of tradable CAFE credits, we model the standard as creating a tradable credit market with credit price t dollars per vehicle-GPM. This means that when an auto manufacturer sells a vehicle with fuel intensity e_j gallons per mile, it must also submit credits valued at te_j for each unit sold.

Auto manufacturing firms produce a choice set of J vehicles, indexed $j \in \{1, ..., J\}$. Marginal production cost is c_j, price is p_j, and fuel intensity in GPM is e_j. In the model, supply is perfectly competitive, so price equals total marginal cost: $p_j = p_j(t) = c_j + te_j$.[12] Like some prior literature, we assume that the choice set is fixed, so automakers comply with fuel economy standards by increasing the relative price of low-MPG vehicles, instead of by introducing more hybrid vehicles or MPG-improving technologies.

Consumers choose exactly one option from the J vehicles or an outside option indexed $j = 0$. There are L consumer types, each with different preferences; l indexes types and i indexes consumers within a type. We normalize each consumer type to have measure one consumer. G_{lj} is the present discounted value of fuel cost for vehicle j given fuel intensity e_j and consumer type l's utilization patterns. Consumer i of type l who buys vehicle j enjoys true utility $U_{ilj} = \eta_l(Z_l - p_j - G_{lj}) + \xi_{lj} + \epsilon_{ij}$, where Z_l is income, ξ_{lj} is utility from vehicle use (i.e. utility from vehicle attributes other than price and fuel cost), and ϵ_{ij} is a logit taste shock. Notice that although we assume a particular functional form over ϵ_{ij} to simplify the derivations, preferences are very general because η_l, G_{lj}, and ξ_{lj} can vary arbitrarily across types.

Consumers are potentially biased: when choosing a vehicle, imperfect information or inattention cause them to perceive fuel costs $(1 + b_{lj})G_{lj}$ instead of G_{lj}. Their vehicle choices thus maximize decision utility $\tilde{U}_{ilj} = \eta_l(Z_l - p_j - (1+b_{lj})G_{lj}) + \xi_{lj} + \epsilon_{ij}$. $b_{lj} = 0$ implies no bias. Positive b_{lj} means that the consumer overestimates fuel costs and thus would get more utility than expected because there is additional money left to buy more units of the numeraire good. Conversely, negative b_{lj} means that the consumer underestimates fuel costs and thus would get less utility than expected. Define $\mathbf{b}_l$ as type l's vector of biases for each of the J vehicles.

Given decision utility $\tilde{U}_{ilj}$, the representative decision utility and choice probabilities are standard for the logit model. For any credit price t and any bias b_j, representative decision utility is $V_{lj}(t, b_j) = \eta_l (Z_l - p_j(t) - (1 + b_{lj})G_{lj}) + \xi_{lj}$, and the logit choice probability for any vector of biases $\mathbf{b}$ is $P_{lj}(t, \mathbf{b}) = \frac{\exp(V_{lj}(t,b_j))}{\sum_k \exp(V_{lk}(t,b_k))}$, where j and k both index vehicles.

The aggregate value of fuel economy credit revenues is $T(t) = \sum_l \sum_j te_j P_{lj}(t, \mathbf{b}_l)$. If credits must be bought from the government, we assume that these revenues are recycled to consumers in

[12] In reality, the vehicle market is of course not perfectly competitive. The propositions below also hold with markups that are non-zero but identical across vehicles. When markups vary across vehicles, the optimal fuel economy standard also depends on the covariance between markup and fuel economy, and the optimal policy formula has an additional term reflecting this. If, as is likely to be the case, markups are higher for low-fuel economy vehicles, then the optimal standard is less stringent than under perfect competition.

lump-sum payments. If credits are grandfathered to auto manufacturers, as is essentially the case under the current policy, then these revenues enter as producer surplus.

We define the "stringency" of the fuel economy standard as $S(t) \equiv \sum_l \sum_j e_j \left[P_{lj}(t, \mathbf{b}_l) - P_{lj}(0, \mathbf{b}_l)\right]$. In words, S is the required change in sales-weighted average fuel intensity relative to the baseline with no standard. $S < 0$ reflects a decrease in fuel intensity, i.e. an increase in fuel economy. Because higher t increases the relative price of higher-fuel intensity vehicles, there is a unique and monotonically decreasing relationship between S and t: the more stringent the required fuel intensity reduction, the higher the credit price. The policymaker sets t (or equivalently, S) to maximize social welfare, which is the sum of true utility across consumer types:

$$W(t) = \underbrace{T(t)}_{Credit\ revenue} + \sum_l \left[\underbrace{\frac{1}{\eta_l} \ln \left(\sum_j \exp\left(V_{lj}(t, b_{lj})\right) \right)}_{Perceived\ consumer\ surplus} + \underbrace{\sum_j b_{lj} G_{lj} P_{lj}(t, \mathbf{b}_l)}_{Bias} \right]. \tag{4}$$

The first terms $T(t)$ reflects credit revenues. The second term is perceived consumer surplus, from the standard Small and Rosen (1981) formula. The final term is the bias: the expected difference $b_{lj}G_{lj}$ between perceived and true consumer surplus, summing over vehicles and weighting by choice probability P_{lj}.

Ideally, the policymaker could achieve the first best by providing what we call a "pure nudge" – that is, an information provision intervention that removes bias, causing all consumers to now have $\mathbf{b}_l = \mathbf{0}$. Alternatively, the first best would obtain under a hypothetical system of type-by-vehicle-level taxes that exactly offset each type's bias in evaluating each vehicle: $\tau^*_{lj} = -b_{lj}G_{lj}$. Of course, such individually-tailored taxes are not practical. A "pure nudge" is probably also unrealistic, and would certainly be costly: our information provision intervention took a meaningful amount of consumers' time to deliver, and it only provided information about a few vehicles. For this reason, the social planner is constrained to considering the second-best social optimum under a fuel economy standard.

V.B Results

We use this framework to show two propositions that illustrate the potential policy implications of our empirical estimates. The first proposition extends the optimal tax formula of Allcott and Taubinsky (2015) to this multi-product setting.

Proposition 1. The optimal fuel economy standard imposes a credit price equal to the average marginal bias:

$$t^* = \frac{-\sum_l \sum_j \frac{dP_{lj}}{dt} b_{lj} G_{lj}}{\sum_l \sum_j \frac{dP_{lj}}{dt} e_{lj}}. \tag{5}$$

The numerator is the average bias (in dollar terms), weighted by the demand slopes. The result that the optimal internality tax equals the average marginal internality parallels the Diamond (1973) result that the optimal externality tax equals the average marginal externality. The denominator translates this average marginal bias from units of dollars to units of dollars per unit fuel intensity.

To see this most clearly, imagine that all consumers undervalue fuel costs by the same proportion, so $b_{lj} = b < 0$. Further imagine that $G_{lj} = \chi e_j$, where χ reflects discount rates and driving patterns and is constant across consumers. Then the optimal credit price is just $t^* = -b\chi$ per unit of fuel intensity, i.e. a tax that exactly offsets the bias in evaluating each vehicle.

For the second proposition, define Q as the effect of a pure nudge on sales-weighted average fuel intensity: $Q \equiv \sum_l \sum_j e_j \left[P_{lj}(0, \mathbf{0}) - P_{lj}(0, \mathbf{b}_l)\right]$.

Proposition 2: If b and χ are homogeneous, the socially-optimal fuel economy standard reduces fuel intensity by the same amount as a pure nudge:

$$S(t^*) = Q. \tag{6}$$

Proposition 2 is a crucial result for interpreting our treatment effect estimates. If our informational interventions can be interpreted as a pure nudge that addresses imperfect information and inattention, then our estimated effects on average fuel intensity from Table 4 are estimates of Q.

See Appendix C for proofs. The Appendix also shows the Proposition 2 holds even if b and χ are heterogeneous, as long as the "mistargeting" of the fuel economy standard – that is, the difference between the CAFE credit cost and the bias in evaluating each vehicle – is orthogonal to fuel intensity and true preferences across vehicles.

V.C Interpreting Empirical Results

If fuel economy information provision can be interpreted as a pure nudge, the average treatment effect of information on vehicle fuel intensity equals Q. Thus, per Proposition 2, the average treatment effects estimated in Section IV equal the fuel economy standard that would be justified by imperfect information and inattention.

There are several reasons to be very cautious about interpreting our informational interventions as pure nudges. First, even if our intervention provided clear information about the consideration set, it did not present information about the rest of the choice set. Second, the treatment groups may have forgotten the information provided by the time they actually decided what vehicle to purchase, as evidenced by the decay of stated preference treatment effects between Tables 2 and 3.

Third, not everyone in our treatment groups may have paid attention to the information we provided. To help mitigate this, we had the dealership RAs record whether people had completed the intervention, and we required online experiment participants to answer quizzes before completing the intervention.

Fourth, the interventions could have induced experimenter demand effects, in which participants changed their vehicle purchases to conform to what they perceived the researchers wanted. To address this, we clearly communicated to the dealership RAs that "our explicit goal is not to influence consumers to pursue fuel-efficient vehicles. Rather, we are exploring the ways in which the presentation of information affects ultimate purchasing behavior." In any event, we believe that it is unlikely that experimenter demand effects would meaningfully influence such large purchases, especially given that experiment participants typically did not make purchases the same day as the intervention and were likely uncertain as to whether they would ever hear from us again. Any experimenter demand effects would likely increase the treatment effects, which biases against our result of zero effect.

Even if our intervention is a pure nudge, our sample are not representative of the U.S. population, both because of selection into the original randomized sample and attrition from that sample to the final sample for which we have data. To help mitigate this issue, we ran two experiments in very different populations and reweighted on observables. Of course, both of our samples likely still differ in unobservable ways from the policy-relevant target population.

While keeping these concerns in mind, Table 6 details how our treatment effects could be interpreted in the context of Proposition 2. The top panel presents approximations of stringency $S(t)$ for the CAFE standards currently in effect. The objective of the "counterfactual" is to establish the average fuel intensity that would arise but for CAFE standards, or $\sum_l \sum_j e_j P_{lj}(0, \mathbf{b}_l)$ in our model. The appropriate counterfactual depends on assumptions about technological change and consumer demand, and gas prices in particular. As a simple benchmark, we use the sales-weighted average fuel economy for model year 2005 vehicles, when gas prices were very similar to their current (2016 average) levels and the new standards had not yet been promulgated. This may be too high, as CAFE standards were already binding for some automakers in 2005, or too low, as technological change and consumer preferences could have evolved since then in the absence of the regulation.

We calculate stringency of the CAFE regulation as of 2016 and 2025 by subtracting the regulatory requirement in each year from the 2004 counterfactual. For the 2016 regulatory requirement, we directly use sales-weighted fuel economy of model year 2016 vehicles from EPA (2016). For 2025, we use the fuel economy that the NHTSA (2012) projects will be achieved, after accounting for various alternative compliance strategies. Subtracting the counterfactuals from the regulatory requirements gives fuel intensity decreases of 1.12 and 2.26 gallons per hundred miles in 2016 and 2025, respectively, or increases of 5.7 and 16.2 MPG.

The bottom panel recaps our key treatment effect estimates from Section IV. Column 1 is re-stated directly from previous tables, while the results in units of MPG in column 2 are from re-estimating the same regressions with fuel economy in MPG as the dependent variable. The stated preference results from Panel (c) of Table 2 would justify a required decrease of 0.032 gallons per 100 miles, or equivalently an increase of 0.19 MPG. The stated preference results address the first two concerns listed above, by considering a choice that was only between vehicles about which we had provided information and that was made immediately after the information was provided.

The revealed preference estimates from Table 4 show statistically zero effect. The 90 percent confidence intervals for the dealership and online experiments, respectively, reject fuel intensity decreases of more than -0.09 and -0.04 gallons per 100 miles in sample, and -0.50 and -0.08 when re-weighted for national representativeness on observables. When re-estimated with the dependent variable in MPG, the confidence bounds for the two experiments are 1.24 and 0.28 MPG, respectively, or 3.28 and 0.62 MPG when re-weighted.

Thus, depending on which experiment and weights we use, the current and proposed CAFE standards are perhaps an order of magnitude larger than the largest likely treatment effects of information. This means that even though our interventions are unlikely to be a "pure nudge," and even though our samples are unlikely to be representative, the true effects of a pure nudge in the national population would have to be *dramatically* different than our estimates to be valid as a significant justification for the current CAFE standards.

Table 6: **Treatment Effects vs. Actual CAFE Standards**

	(1) Gallons per 100 miles	(2) Miles per gallon
Current CAFE Standards		
"Counterfactual" (2005 sales)	5.03	19.9
2016 sales	3.91	25.6
2025 CAFE standard	2.77	36.1
"2016 stringency": 2016 sales – Counterfactual	-1.12	5.7
"2025 Stringency": 2025 CAFE standard – Counterfactual	-2.26	16.2
Treatment Effects of Information		
Stated preference (point estimate; Table 2, Panel (c))	-0.03	0.19
Revealed preference (90% confidence bound; Table 4)		
Dealership experiment, equally-weighted (column 2)	-0.09	1.24
Dealership experiment, re-weighted (column 3)	-0.50	3.28
Online experiment, equally-weighted (column 5)	-0.04	0.28
Online experiment, re-weighted (column 6)	-0.08	0.62

Notes: The top panel details the CAFE standards currently in effect for light-duty vehicles. Sales-weighted adjusted fuel economy for model years 2005 and 2016 are from Table 2.1 of EPA (2016). The 2025 CAFE standard is the "achieved" unadjusted sales-weighted MPG of 46.2 from NHTSA (2012), multiplied by 0.782 to transform to adjusted MPG; the 0.782 adjustment factor reflects data for the most recent year in Table 10.1 of EPA (2016). The bottom panel presents the treatment effects of information, as estimated in Tables 2 and 4. In the bottom panel, the miles per gallon estimates in column 2 are calculated by re-estimating Equation (3) with fuel economy in miles per gallon as the dependent variable.

Our model does not include externalities or other justifications for CAFE other than informational and attentional biases. Thus, our analysis can be viewed as evaluating these biases in isolation as a justification for CAFE. This is still relevant, because as described in the introduction, the Regulatory Impact Analyses rely largely on consumers' private net benefits – not externalities – to justify the stringency of the policy.[13]

VI Conclusion

It has long been argued that car buyers are poorly-informed, inattentive, or otherwise cognitively-constrained when evaluating fuel economy, and that this causes them to buy systematically lower-fuel economy vehicles than would be optimal. We tested this hypothesis with two information provision field experiments. In both experiments, we find that our treatments did not have a statistically or economically significant effect on the fuel economy of purchased vehicles.

[13] Allcott, Mullainathan, and Taubinsky (2014) present a model that includes externalities, as well as other extensions such as a vehicle utilization margin (the decision of how much to drive) and gas taxes as a potential policy instrument.

There are perhaps two main interpretations of these results. The first is that while our interventions did draw attention to fuel economy for a few minutes, the information we provided was not very useful, and/or people soon forgot it. Put simply, the interventions did not come close to fully informing people about fuel economy. This would be deeply interesting, as it would highlight the difficulties in providing product information. New cars already have fuel economy information labels prominently posted in the windows, and the Environmental Protection Agency has a useful fuel economy information website, www.fueleconomy.gov. Then, in addition, our dealership intervention provided in-person, individually-tailored fuel economy information via a well-designed iPad app. If, after all of these efforts, we still need stringent fuel economy standards to address lack of information about fuel economy, this is a striking testament to the deficiencies in currently-feasible information provision technologies.

The second interpretation is to take the empirical estimates more seriously in the context of our optimal policy model, suggesting that imperfect information, inattention, and related cognitive constraints do not have a significant systematic effect on vehicle markets. This would imply either that some other market failure or behavioral failure must justify the CAFE standard, or that the net private benefits projected in the CAFE Regulatory Impact Analyses do not actually exist. The latter possibility would arise if the RIAs' engineering models did not account for the full fixed costs, production costs, or performance reductions from fuel economy-improving technologies. In this case, there would still be an economic justification for fuel economy standards as a second-best externality policy – albeit a highly inefficient one, as shown by Jacobsen (2013). But if fuel economy is more expensive than the RIA models assume, the socially-optimal CAFE standard would likely be significantly less stringent than the current or proposed levels.

References

[1] Abaluck, Jason, and Jonathan Gruber (2011). "Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *American Economic Review* 101 (4):1180-1210.

[2] Allcott, Hunt (2011). "Consumers' Perceptions and Misperceptions of Energy Costs." *American Economic Review* 101 (3): 98-104

[3] Allcott, Hunt (2013). "The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market." *American Economic Journal: Economic Policy* 5 (3): 30-66.

[4] Allcott, Hunt (2015). "Site Selection Bias in Program Evaluation." *Quarterly Journal of Economics* 130 (3): 1117-1165.

[5] Allcott, Hunt, Sendhil Mullainathan, and Dmitry Taubinsky (2014). "Energy Policy with Externalities and Internalities." *Journal of Public Economics* 112: 72-88.

[6] Allcott, Hunt, and Richard L. Sweeney (2017). "The Role of Sales Agents in Information Disclosure: Evidence from a Field Experiment." *Management Science* forthcoming.

[7] Allcott, Hunt, and Dmitry Taubinsky (2015). "Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market." *American Economic Review* 105 (8): 2501-2538.

[8] Allcott, Hunt, and Nathan Wozny (2014). "Gasoline Prices, Fuel Economy, and the Energy Paradox." *Review of Economics and Statistics* 96 (10): 779-795.

[9] Austin, David, and Terry Dinan (2005). "Clearing the Air: The Costs and Consequences of Higher CAFE Standards and Increased Gasoline Taxes." *Journal of Environmental Economics and Management* 50 (3): 562-582.

[10] Barber, Brad M., Terrance Odean, and Lu Zheng (2005). "Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows." *Journal of Business* 78 (6): 2095-2120.

[11] Bhargava, Saurabh, and Dayanand Manoli (2015). "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment." *American Economic Review* 105 (11): 3489-3529.

[12] Bollinger, Brian, Phillip Leslie, and Alan Sorensen (2011). "Calorie Posting in Chain Restaurants." *American Economic Journal: Economic Policy* 3 (1): 91-128.

[13] Busse, Meghan, Christopher Knittel, and Florian Zettelmeyer (2013). "Are Consumers Myopic? Evidence from New and Used Car Purchases." *American Economic Review* 103 (1): 220-256.

[14] Chetty, Raj, Adam Looney, and Kory Kroft (2009). "Salience and Taxation: Theory and Evidence." *American Economic Review* 99 (4): 1145-1177.

[15] Choi, James J., David Laibson, and Brigitte C. Madrian (2010). "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds." *Review of Financial Studies* 23 (4): 1405–1432.

[16] Davis, Lucas W., and Gilbert E. Metcalf (2016). "Does Better Information Lead to Better Choices? Evidence from Energy-Efficiency Labels." *Journal of the Association of Environmental and Resource Economists*, 3 (3): 589-625.

[17] Diamond, Peter (1973). "Consumption Externalities and Imperfect Corrective Pricing." *Bell Journal of Economics and Management Science* 4 (2): 526-538.

[18] Dolan, Paul, and Robert Metcalfe (2013). "Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation." CEP Discussion Papers DP1222, Centre for Economic Performance, LSE.

[19] Dranove, David, and Ginger Zhe Jin (2010). "Quality Disclosure and Certification: Theory and Practice." *Journal of Economic Literature* 48 (4): 935-963.

[20] Dubin, Jeffrey, and Daniel McFadden (1984). "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption." *Econometrica* 52 (2): 345-362.

[21] Duflo, Esther, and Emmanel Saez (2003). "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics* 118 (3): 815–842.

[22] Dupas, Pascaline (2011). "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya." *American Economic Journal: Applied Economics* 3 (1): 1-34.

[23] EPA (U.S. Environmental Protection Agency) (2010). "Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards; Final Rule." *Federal Register* 75 (88): 25324-25728.

[24] EPA (U.S. Environmental Protection Agency) (2012). "Regulatory Impact Analysis: Final Rulemaking for 2017-2025 Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards." EPA-420-R-12-016 (August).

[25] EPA (U.S. Environmental Protection Agency) (2016). "Light-Duty Automotive Technology, Carbon Dioxide Emission, and Fuel Economy Trends: 1975 Through 2016." EPA-420-R-16-010 (November).

[26] Farhi, Emmanuel, and Xavier Gabaix (2015). "Optimal Taxation with Behavioral Agents." Working Paper, New York University (August).

[27] Ferraro, Paul J., and Michael K. Price (2013). "Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment." *Review of Economics and Statistics* 95 (1): 64-73.

[28] Gabaix, Xavier, and David Laibson (2006). "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics* 121 (2): 505-540.

[29] Gayer, Ted (2011). "A Better Approach to Environmental Regulation: Getting the Costs and Benefits Right." Hamilton Project Discussion Paper 2011-06 (May).

[30] Goldberg, Pinelopi (1998). "The Effects of the Corporate Average Fuel Economy Standards in the US." *Journal of Industrial Economics* 46: 1-33.

[31] Greene, David (2010). "How Consumers Value Fuel Economy: A Literature Review." US Environmental Protection Agency Technical Report EPA-420-R-10-008. Washington, DC, March.

[32] Greene, David L., Philip D. Patterson, Margaret Singh, and Jia Li (2005). "Feebates, Rebates and Gas-guzzler Taxes: A Study of Incentives for Increased Fuel Economy." *Energy Policy* 33: 757-775.

[33] Grigolon, Laura, Mathias Reynaert, and Frank Verboven (2015). "Consumer Valuation of Fuel Costs and the Effectiveness of Tax Policy: Evidence from the European Car Market." Working Paper, University of Leuven (July).

[34] Grubb, Michael (2009). "Selling to Overconfident Consumers." *American Economic Review* 99 (5): 1770-1807.

[35] Gruber, Jonathan, and Botond Koszegi (2004). "Tax Incidence when Individuals are Time-Inconsistent: The Case of Cigarette Excise Taxes." *Journal of Public Economics* 88: 1959-1987.

[36] Hainmueller, Jens (2012). "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." *Political Analysis* 20: 25-46.

[37] Handel, Benjamin R., and Jonathan T. Kolstad (2015). "Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare." *American Economic Review* 105 (8): 2449-2500.

[38] Hastings, Justine S., and Jeffrey M. Weinstein (2008). "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." *Quarterly Journal of Economics* 123 (4): 1373-1414.

[39] Hausman, Jerry (1979). "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables." *Bell Journal of Economics* 10 (1): 33-54.

[40] Heutel, Garth (2015). "Optimal Policy Instruments for Externality-Producing Durable Goods under Present Bias." *Journal of Environmental Economics and Management* 72: 54-70.

[41] Hossain, Tanjim, and John Morgan (2006). "...Plus Shipping and Handling: Revenue (Non)Equivalence in Field Experiments on eBay." *Advances in Economic Analysis and Policy* 6.

[42] Howell, Sabrina T. (2016). "Joint Ventures and Technology Adoption: A Chinese Industrial Policy that Backfired." Working Paper, New York University.

[43] Ito, Koichiro, and James M. Sallee (2014). "The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards." NBER Working Paper No. 20500 (September).

[44] Jacobsen, Mark R (2013). "Evaluating US Fuel Economy Standards in a Model with Producer and Household Heterogeneity." *American Economic Journal: Economic Policy* 5 (2): 148-187.

[45] Jacobsen, Mark R., and Arthur A. Van Benthem (2015). "Vehicle Scrappage and Gasoline policy." *American Economic Review* 105 (3): 1312-1338.

[46] Jensen, Robert (2010). "The Perceived Returns to Education and the Demand for Schooling." *Quarterly Journal of Economics* 125 (2): 515-548.

[47] Jessoe, Katrina, and David Rapson (2015). "Commercial and Industrial Demand Response Under Mandatory Time-of-Use Electricity Pricing." *Journal of Industrial Economics* 63 (3): 397-421.

[48] Jin, Ginger Zhe, and Alan T. Sorensen (2006). "Information and Consumer Choice: The Value of Publicized Health Plan Ratings." *Journal of Health Economics* 25 (2): 248-275.

[49] Kempton, Willett, and Laura Montgomery (1982). "Folk Quantification of Energy." *Energy* 7 (10): 817-827.

[50] Kling, Jeffrey, Sendhil Mullainathan, Eldar Shafir, Lee Vermeulen, and Marian Wrobel (2012). "Comparison Friction: Experimental Evidence from Medicare Drug Plans." *Quarterly Journal of Economics* 127 (1): 199-235.

[51] Larrick, Richard P., and Jack B. Soll (2008). "The MPG Illusion." *Science* 320 (5883): 1593-1594.

[52] Lockwood, Benjamin B., and Dmitry Taubinsky (2017). "Regressive Sin Taxes." Working Paper, Dartmouth College.

[53] Mullainathan, Sendhil, Joshua Schwartzstein, and William Congdon (2012). "A Reduced-Form Approach to Behavioral Public Finance." *Annual Review of Economics* 4: 17.1-17.30.

[54] Newell, Richard G., and Juha Siikamaki (2014). "Nudging Energy Efficiency Behavior: The Role of Information Labels." *Journal of the Association of Environmental and Resource Economists* 1 (4): 555-598.

[55] NHTSA (National Highway Traffic Safety Administration) (2012). "Final Regulatory Impact Analysis: Corporate Average Fuel Economy for MY 2017-MY 2025 Passenger Cars and Light Trucks." Office of Regulatory Analysis and Evaluation, National Center for Statistics and Analysis (March).

[56] O'Donoghue, Edward, and Matthew Rabin (2006). "Optimal Sin Taxes." *Journal of Public Economics* 90: 1825-1849.

[57] Reynaert, Mathias, and James M. Sallee (2016). "Corrective Policy and Goodhart's Law: The Case of Carbon Emissions from Automobiles." NBER Working Paper No. 22911 (December).

[58] Sallee, James M., Sarah E. West, and Wei Fan (2016). "Do Consumers Recognize the value of Fuel Economy? Evidence from Used Car Prices and Gasoline Price Fluctuations." *Journal of Public Economics* 135: 61-73.

[59] Sanstad, Alan, and Richard Howarth (1994). "'Normal' Markets, Market Imperfections, and Energy Efficiency." *Energy Policy* 22 (10): 811-818.

[60] Scanlon, Dennis P., Michael Chernew, Catherine McLaughlin, and Gary Solon (2002). "The Impact of Health Plan Report Cards on Managed Care Enrollment." *Journal of Health Economics* 21 (1): 19-41.

[61] Small, Kenneth A., and Harvey S. Rosen (1981). "Applied Welfare Economics with Discrete Choice Models." *Econometrica* 49 (1): 105-130.

[62] Stern, Paul C., and Elliot Aronson (1984). Energy Use: The Human Dimension. New York: W.H. Freeman and Co.

[63] Turrentine, Thomas, and Kenneth Kurani (2007). "Car Buyers and Fuel Economy?" *Energy Policy* 35: 1213-1223.

[64] Tsvetanov, Tsvetan, and Kathleen Segerson (2013). "Re-evaluating the Role of Energy Efficiency Standards: A Behavioral Economics Approach." *Journal of Environmental Economics and Management* 66 (2): 347-363.

Online Appendix: Not for Publication

Are Consumers Poorly-Informed about Fuel Economy? Evidence from Two Experiments

Hunt Allcott and Christopher Knittel

A Data Appendix

A.A Dealership and Online Survey Data

Basic data cleaning steps for dealership data included the following:

- Some survey observations were test cases. We removed these from the iPad data by inspecting comments by RAs or respondent names for words such as "test" or "fake."
- The follow-up phone survey was delivered twice to some households. In these cases, we kept the more complete observation, or if both were equally complete, one of the repeated observations was randomly chosen.
- Some people provided a range of numbers for expected fuel costs on the follow-up phone survey. In these cases, we used the midpoint of the range.

In the follow-up surveys for both experiments, some people reported a new vehicle purchased that had the same make, model, and model year as their current vehicle in the baseline survey; these cases were coded as not having purchased new cars.

There are a limited number of apparently-careless survey responses, in particular for the stated preference results for the online survey the fuel cost belief data from both surveys. We cleaned these in the following ways:

- We dropped all gasoline price expectations of less than $1 or greater than or equal to $10 per gallon.
- We dropped all expected annual miles driven less than 1,000 or greater than 75,000.
- We dropped all expected vehicle annual fuel costs less than $100 if the respondent reported expecting to drive 2,000 or more miles per year.
- We dropped several common patterns of careless responses, for example writing that annual maintenance, insurance, and fuel costs would all equal $X per year, with X\leq$10.

A.B Fuel Economy, Census, and National Household Travel Survey Data

We use the official EPA vehicle-level fuel economy data available from www.fueleconomy.gov/feg/download.shtml. Vehicles reported in the survey were matched to vehicles in the EPA data based on manufacturer, year, and model name as well as secondary characteristics such as fuel type, transmission, engine size and number of cylinders. If one or more of the secondary characteristics were missing, creating possible matches to more than one vehicle in the EPA data, we used the average fuel economy rating of all such possible matches.

At baseline, individuals report miles they expect to drive and the proportion of city vs. highway driving. Combining these self-reported city/highway proportions with fuel economy numbers from the EPA data, we computed average fuel economy and fuel intensity (defined as inverse of fuel economy) for each person-car combination in the data.

We gathered median income and median education for each respondent's zip code from the 2014 American Community Survey (ACS) 5-year estimates. Mean imputation was used to impute missing values of these and other covariates used in the regressions.

National average covariates in Table 1 were estimated from the 2009 National Household Transportation Survey (NHTS). We define a new car buyer as a household having bought a vehicle with model year 2008 or 2009. Individuals less than 22 years old were dropped while calculating the average household age for it to be closer to that of the household head's. Annual miles driven are from the BESTMILE variable. The NHTS reports "unadjusted" combined fuel economy, which we adjusted using the scaling factors in Table 10.1 of EPA (2016).

B Treatment Effects on Beliefs, and Beliefs as a Moderator

Does the treatment make fuel cost beliefs more precise? And is the treatment effect larger for consumers whose beliefs were more biased at baseline? Appendix Table A1 explores these questions using the online experiment data. We cannot do parallel analyses for the dealership experiment because we did not elicit control group baseline beliefs.[14]

Column 1 first tests whether the treatment reduces the extent to which people systematically over- or understate the fuel cost differences between vehicles. To do this, we limit the sample to those who correctly know which of their first- vs. second-choice vehicle has higher fuel economy, i.e. those with $\phi_i > 0$ at baseline and follow-up. Using that sample, we regress the follow-up valuation ratio on the baseline valuation ratio, the treatment indicator, and the interaction thereof. Results in column 1 show that estimates are very imprecise: we cannot reject that the treatment more than doubles, or fully reverses, the correlation between baseline and follow-up beliefs. Column 2 then tests whether the treatment reduces the amount of noise in people's reported beliefs, presenting a regression of follow-up belief noise $|\phi_i - 1|$ on baseline belief noise and the treatment indicator. Here again, the estimates are imprecise, and we cannot reject that the intervention has a large positive or negative effect on the correlation between baseline and follow-up belief noise.

Columns 3 and 4 present comparable regressions, except with purchased vehicle fuel intensity as the dependent variable. Here again, we cannot reject large possible effects of the treatment relative

[14] We did not want to meaningfully draw attention to fuel costs in the control group. Because the online survey could involve more questions, we asked the above question to both treatment and control, but obscured the importance of fuel costs by also asking parallel questions about insurance and maintenance. Because customers were more hurried in the dealerships, such additional questions were not practical, so we elicited fuel cost beliefs from the treatment group only, at the beginning of the treatment intervention.

to the baseline correlation. Thus, it is not possible to infer whether the treatment makes fuel cost beliefs meaningfully more precise or meaningfully moderates the treatment effect.

Table A1: **Effects on Beliefs, and Beliefs as a Moderator**

	(1) Valuation ratio: purchased - 2nd choice	(2) Abs. belief error: purchased - 2nd choice	(3) Purchased vehicle fuel intensity	(4) Purchased vehicle fuel intensity
Treatment × valuation ratio: 1st - 2nd choice	0.13		0.01	
	(0.09)		(0.06)	
Treatment × abs. belief error: 1st - 2nd choice				-0.05
				(0.06)
Treatment	-0.08	-0.03	0.05	0.11
	(0.10)	(0.07)	(0.07)	(0.09)
Valuation ratio: 1st - 2nd choice	0.06		-0.07	
	(0.07)		(0.05)	
Abs. belief error: 1st - 2nd choice		0.09*		0.03
		(0.04)		(0.05)
N	922	1,181	1,229	1,343
R^2	0.02	0.01	0.01	0.01
Dependent variable mean	0.88	1.33	4.07	4.07

Notes: Columns 1 and exclude observations with negative valuation ratios. The dependent variable in columns 3 and 4 is purchased vehicle fuel intensity (in gallons per 100 miles). Columns 3 and 4 control for gender, age, a Caucasian indicator, natural log of income, miles driven per year, an indicator for whether the current vehicle is a Ford, current vehicle fuel intensity, consideration set average fuel intensity, and treatment group closure time indicators. Robust standard errors are in parentheses. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

C Proofs of Propositions 1 and 2

C.A Proof of Proposition 1

A necessary condition for the socially-optimal credit price t^* is that $\frac{dW(t)}{dt} = 0$. Proposition 1 is derived from this first-order condition, where

$$\begin{aligned}\frac{dW(t)}{dt} \quad = \quad & \underbrace{\sum_l \sum_j \left[\frac{dP_{lj}(t,\mathbf{b}_l)}{dt} te_j + e_j P_{lj}(t,\mathbf{b}_l)\right]}_{\textit{Change in credit revenue}} \\ & - \underbrace{\sum_l \sum_j e_j P_{lj}(t,\mathbf{b}_l)}_{\textit{Change in perceived CS}} \\ & + \underbrace{\sum_l \sum_j b_{lj} G_{lj} \frac{dP_{lj}(t,\mathbf{b}_l)}{dt}}_{\textit{Change in bias}}.\end{aligned} \tag{7}$$

Re-arranging gives

$$t \cdot \sum_l \sum_j \frac{dP_{lj}(t,\mathbf{b}_l)}{dt} e_j = -\sum_l \sum_j b_{lj} G_{lj} \frac{dP_{lj}(t,\mathbf{b}_l)}{dt},$$

and re-arranging further gives Equation (5).

C.B Proof of Proposition 2

In the text, we defined the effect of a pure nudge $Q \equiv \sum_l \sum_j e_j \left[P_{lj}(0,\mathbf{0}) - P_{lj}(0,\mathbf{b}_l)\right]$ and the stringency of the fuel economy standard $S(t) \equiv \sum_l \sum_j e_j \left[P_{lj}(t,\mathbf{b}_l) - P_{lj}(0,\mathbf{b}_l)\right]$. Further define $\Lambda_{lj} \equiv \exp\left(\eta_l(-e_j t^* - b_{lj} G_{lj})\right)$ for all vehicles ($j \geq 1$), and $\Lambda_{l0} = 0$ for the outside option ($j = 0$). Intuitively, Λ_{lj} is the "mistargeting" of the second-best policy: the value (in exponentiated utils) of the distortion between the credit price for vehicle j, which is $e_j t^*$, and the bias that it is intended to offset, which is $b_{lj} G_{lj}$.

If b and χ are homogeneous, then $t^* = -b\chi$, so $-e_j t^* - b_{lj} G_{lj} = e_j b\chi - b\chi e_j = 0$, and thus $\Lambda_{lj} = 1$. (Intuitively, when bias (in dollar terms) is homogeneous, a fuel economy standard that imposes a uniform credit price has no mistargeting.) Therefore,

$$\sum_l \sum_j e_j P_{lj}(t,\mathbf{b}_l) = \sum_l \frac{\sum_j e_j \exp(V_{lj}(0,\mathbf{0})) \cdot \Lambda_{lj}}{\sum_j \exp(V_{lj}(0,\mathbf{0})) \cdot \Lambda_{lj}} = \sum_l \frac{\sum_j e_j \exp(V_{lj}(0,\mathbf{0}))}{\sum_j \exp(V_{lj}(0,\mathbf{0}))} = \sum_l \sum_j e_j P_{lj}(0,\mathbf{0}) \tag{8}$$

We thus have $S(t^*) = \sum_l \sum_j e_j \left[P_{lj}(t, \mathbf{b}_l) - P_{lj}(0, \mathbf{b}_l)\right] = \sum_l \sum_j e_j \left[P_{lj}(0, \mathbf{0}) - P_{lj}(0, \mathbf{b}_l)\right] = Q$.

Proposition 2 also holds if the following orthogonality conditions hold across all vehicles j, within all types l: $Cov\left(e_j \exp(V_{lj}(0, \mathbf{0})), \Lambda_{lj}\right) = 0$ and $Cov\left(\exp(V_{lj}(0, \mathbf{0}), \Lambda_{lj}\right) = 0$. Intuitively, these conditions require that the mistargeting of the second best policy Λ_{lj} is unrelated to fuel intensity e_j and true preferences $V_{lj}(0, \mathbf{0})$. Under these conditions, the second equality in Equation (8) holds because

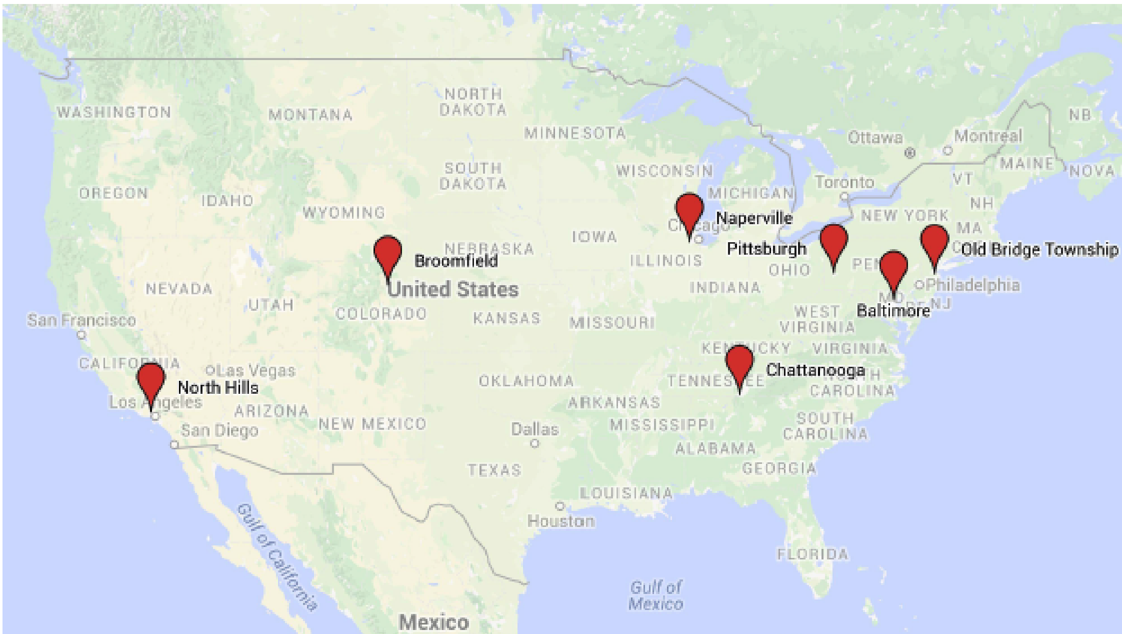
$$\sum_l \frac{\sum_j e_j \exp(V_{lj}(0,\mathbf{0})) \cdot \Lambda_{lj}}{\sum_j \exp(V_{lj}(0,\mathbf{0})) \cdot \Lambda_{lj}} = \sum_l \frac{\left[\sum_j e_j \exp(V_{lj}(0,\mathbf{0}))\right] \cdot \left[\sum_j \Lambda_{lj}\right] + J^2 Cov\left(e_j \exp(V_{lj}(0,\mathbf{0})), \Lambda_{lj}\right)}{\left[\sum_j \exp(V_{lj}(0,\mathbf{0}))\right] \cdot \left[\sum_j \Lambda_{lj}\right] + J^2 Cov\left(\exp(V_{lj}(0,\mathbf{0}), \Lambda_{lj}\right)} \tag{9}$$

$$= \sum_l \frac{\left[\sum_j e_j \exp(V_{lj}(0,\mathbf{0}))\right] \cdot \left[\sum_j \Lambda_{lj}\right]}{\left[\sum_j \exp(V_{lj}(0,\mathbf{0}))\right] \cdot \left[\sum_j \Lambda_{lj}\right]} = \sum_l \frac{\sum_j e_j \exp(V_{lj}(0,\mathbf{0}))}{\sum_j \exp(V_{lj}(0,\mathbf{0}))}, \tag{10}$$

where the equality between the first and second lines holds due to the orthogonality conditions.

D Appendix Tables and Figures

Figure A1: **Ford Dealership Experiment Locations**



Notes: This map shows the locations of the seven Ford dealerships in the dealership information provision experiment.

Table A2: **Treatment Group Balance on Observables**

	Treatment	Control	Difference
Male	0.57	0.59	-0.01
	(0.01)	(0.01)	(0.02)
Age	40.20	40.02	0.18
	(0.37)	(0.37)	(0.53)
Caucasian	0.69	0.71	-0.02
	(0.01)	(0.01)	(0.02)
Income ($000s)	72.26	73.04	-0.78
	(0.79)	(0.78)	(1.11)
Miles driven/year (000s)	14.64	15.37	-0.72
	(0.36)	(0.48)	(0.61)
Current vehicle is Ford	0.35	0.37	-0.01
	(0.02)	(0.01)	(0.02)
Current fuel intensity (gallons/100 miles)	4.66	4.77	-0.11**
	(0.04)	(0.04)	(0.05)
Consideration set fuel intensity (gallons/100 miles)	4.26	4.38	-0.12**
	(0.04)	(0.04)	(0.05)
p-value of F-test of joint significance		0.18	
N	958	1,031	1,989

(a) **Dealership Experiment**

	Treatment	Control	Difference
Male	0.56	0.57	-0.01
	(0.01)	(0.01)	(0.01)
Age	54.52	54.49	0.03
	(0.23)	(0.27)	(0.36)
Caucasian	0.84	0.83	0.00
	(0.01)	(0.01)	(0.01)
Income ($000s)	110.57	117.49	-6.92**
	(1.83)	(2.89)	(3.26)
Miles driven/year (000s)	11.48	11.54	-0.06
	(0.13)	(0.17)	(0.21)
Current vehicle is Ford	0.12	0.11	0.00
	(0.01)	(0.01)	(0.01)
Current fuel intensity (gallons/100 miles)	4.61	4.61	0.00
	(0.02)	(0.02)	(0.03)
Consideration set fuel intensity (gallons/100 miles)	4.15	4.13	0.03
	(0.01)	(0.02)	(0.02)
p-value of F-test of joint significance		0.25	
N	3,771	2,545	6,316

(b) **Online Experiment**

Notes: These tables present tests of balance between treatment and control groups in the dealership and online experiments. In each case, the sample is the set of observations that were allocated to treatment or control. The bottom row reports the p-value of an F-test of a regression of the treatment indicator on all covariates. Standard errors in parentheses. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Table A3: **Attrition by Treatment Condition**

	(1) Dealership	(2) Online
Treatment	0.004	0.016
	(0.018)	(0.011)
N	1,989	6,316
R^2	0.00	0.02
Dependent variable mean	0.81	0.76

Notes: This table presents regressions of an attrition indicator variable on the treatment indicator variable, in the sample of valid observations that were allocated to treatment or control. Estimates with the online experiment data also include treatment group closure time indicators. Robust standard errors are in parentheses. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Table A4: **Tests of Differential Attrition by Demographics**

	(1) Dealership	(2) Online
Male	-0.057**	-0.053***
	(0.027)	(0.018)
Age	-0.003**	0.000
	(0.001)	(0.001)
Caucasian	-0.051*	-0.021
	(0.030)	(0.023)
ln(Income)	-0.029	-0.038***
	(0.038)	(0.011)
Miles driven/year (000s)	-0.001	-0.000
	(0.001)	(0.001)
Current vehicle is Ford	-0.024	-0.023
	(0.026)	(0.029)
Current fuel intensity (gallons/100 miles)	0.007	0.008
	(0.011)	(0.009)
Consideration set fuel intensity (gallons/100 miles)	0.006	-0.009
	(0.012)	(0.011)
Treatment × Male	0.026	0.027
	(0.040)	(0.023)
Treatment × Age	0.003*	-0.000
	(0.002)	(0.001)
Treatment × Caucasian	-0.024	-0.014
	(0.043)	(0.029)
Treatment × ln(Income)	0.050	0.019
	(0.053)	(0.015)
Treatment × Miles driven/year (000s)	0.000	-0.001
	(0.001)	(0.001)
Treatment × Current vehicle is Ford	0.014	0.022
	(0.039)	(0.036)
Treatment × Current fuel intensity (gallons/100 miles)	-0.001	-0.005
	(0.016)	(0.012)
Treatment × Consideration set fuel intensity (gallons/100 miles)	-0.013	-0.000
	(0.017)	(0.014)
N	1,989	6,316
R^2	0.01	0.03
Dependent variable mean	0.81	0.76

Notes: This table presents regressions of an attrition indicator variable on the treatment indicator variable and interactions with demographic covariates, in the sample of valid observations that were allocated to treatment or control. Estimates with the online experiment data also include treatment group closure time indicators. Robust standard errors are in parentheses. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Table A5: **Are Elicited Beliefs Meaningful?**

	(1) Valuation ratio: purchased	(2) Valuation ratio: purchased - 2nd choice	(3) Purchased vehicle fuel intensity	(4) Abs. belief error: purchased - 2nd choice
Valuation ratio: 1st choice	0.542***			
	(0.128)			
Valuation ratio: 1st - 2nd choice		0.305	0.021	
		(0.190)	(0.122)	
Valuation ratio: purchased - 2nd choice			-0.108	
			(0.098)	
Abs. belief error: 1st - 2nd choice				0.280
				(0.183)
N	126	44	44	58
R^2	0.28	0.07	0.02	0.05
Dependent variable mean	0.96	1.04	4.09	1.75

(a) **Dealership Experiment**

	(1) Valuation ratio: purchased	(2) Valuation ratio: purchased - 2nd choice	(3) Purchased vehicle fuel intensity	(4) Abs. belief error: purchased - 2nd choice
Valuation ratio: 1st choice	0.396***			
	(0.034)			
Valuation ratio: 1st - 2nd choice		0.145***	-0.045	
		(0.045)	(0.034)	
Valuation ratio: purchased - 2nd choice			-0.093***	
			(0.026)	
Abs. belief error: 1st - 2nd choice				0.093**
				(0.047)
N	1,255	922	922	1,126
R^2	0.18	0.01	0.02	0.01
Dependent variable mean	1.07	0.88	4.06	1.33

(b) **Online Experiment**

Notes: In column 1, valuation ratios are the ratio of perceived to actual annual fuel cost, calculated using Equation (1). In columns 2 and 3, valuation ratios are the ratio of perceived to annual fuel cost differences between the two vehicles, calculated using Equation (2). In column 4, the absolute belief error is the absolute value of the valuation ratio (from Equation (2)) minus one. Columns 2 and 3 exclude observations with negative valuation ratios. Robust standard errors are in parentheses. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

Table A6: **Separate Estimates of Effects for Each of the Four Online Treatments**

	(1) Stated preference	(2) Purchased vehicle
Base Only	-0.028***	0.006
	(0.007)	(0.063)
Base + Relative	-0.026***	0.037
	(0.009)	(0.065)
Base + Climate	-0.034***	0.124**
	(0.007)	(0.059)
All	-0.040***	-0.055
	(0.008)	(0.070)
N	5,018	1,489
R^2	0.97	0.39
Dependent variable mean	4.08	4.09
p-value(Treatment effects equal)	0.53	0.12
p-value(Treatment effects equal 0)	0.00	0.16

Notes: This table presents estimates of Equation (3), with separate treatment indicators for each of the four online treatment groups. In column 1, the dependent variable is the weighted average fuel intensity (in gallons per 100 miles) of the two vehicles in the consideration set, weighted by post-intervention stated purchase probability. In column 2, the dependent variable is weighted average fuel intensity of the vehicle the consumer actually purchased, using data from the follow-up survey. Both columns control for gender, age, a Caucasian indicator, natural log of income, miles driven per year, an indicator for whether the current vehicle is a Ford, current vehicle fuel intensity, consideration set average fuel intensity, and treatment group closure time indicators. Robust standard errors are in parentheses. *, **, ***: statistically different from zero with 90, 95, and 99 percent probability, respectively.

MIT CEEPR

MIT Center for Energy and Environmental Policy Research

MIT Center for Energy and Environmental Policy Research
77 Massachusetts Avenue, E19-411
Cambridge, MA 02139
USA

Website: ceepr.mit.edu

MIT CEEPR Working Paper Series is published by the MIT Center for Energy and Environmental Policy Research from submissions by affiliated researchers.

Email ceepr@mit.edu
Phone (617) 253-3551
Fax (617) 253-9845