Estimating Willingness to Pay for E85 in the United States Using an Intercept Survey of Flex Motorists

Kenneth Liao, Sébastien Pouliot, and Bruce Babcock

Working Paper 16-WP 562
Updated June 2018

Kenneth Liao is Assistant Professor, Department of Economics, Farmingdale State College, Farmingdale, NY 11735, E-mail: liaok@farmingdale.edu.

Sébastien Pouliot is Associate Professor, Department of Economics, Iowa State University, Ames, IA 50011. E-mail: pouliot@iastate.edu.

Bruce A. Babcock is Professor, School of Public Policy, and associate director of the Agricultural Issues Center at the University of California, Riverside, CA 95616. E-mail: bruce.babcock@ucr.edu.

For questions or comments about the contents of this paper, please contact Kenneth Liao, kliao@iastate.edu.

Estimating Willingness to Pay for E85 in the United States

Using an Intercept Survey of Flex Motorists

Sébastien Pouliot

Kenneth Liao

and

Bruce A. Babcock

Abstract
Compliance with the Renewable Fuel Standard (RFS) in United States will require price incentives for a substantial number of motorists with flex-fuel vehicles to switch to high ethanol-gasoline blends. Existing estimates of motorists' willingness to pay for high-ethanol blends use data from Brazil, data generated when prices greatly favored low-ethanol blends, or stated preference data collected from mail and online surveys. We conducted an intercept survey of flex motorists as they refueled in five US states. We overcome the problem caused by sample prices favoring low-ethanol blends by augmenting revealed preference data with stated preference data. A sample-selection problem arises because motorists with high willingness to pay seek out the relatively few stations that sell high-ethanol blends. We use responses from two questions to inform sample selection. We find the average US motorist requires a substantial discount to switch to high ethanol blends beyond the price that equates the cost per mile of driving.

Keywords: Biofuel, E85, Ethanol, Gasoline, Renewable Fuel Standard, Intercept Survey.

JEL codes: Q18, Q41, Q42, C25

Running head: Estimating WTP for E85 using an intercept survey

The Renewable Fuel Standard (RFS) in the United States uses biofuel-blending mandates to achieve its policy objectives of reductions in greenhouse gas emissions, reductions in imports of fossil fuels, and enhanced rural incomes. The blending mandates across biofuel categories are set annually by the Environmental Protection Agency (EPA) and must be met by obligated parties, who are motor fuel producers (oil refineries that produce gasoline and diesel) and motor fuel importers.[1]

EPA sets an overall renewable fuel mandate within which separate mandates exist for advanced biofuels. The difference between the overall renewable fuel mandate and the advanced biofuel mandate is the portion of the overall mandate for conventional biofuel, i.e., corn ethanol. Until 2016, consumption of E10, which is a blend petroleum-based gasoline and no more than 10 percent ethanol, was sufficient to meet the corn ethanol mandate. With practically all US gasoline containing 10 percent ethanol by default, understanding consumer preferences about ethanol had less urgency because the cost of meeting EPA blending mandates did not depend on inducing consumers to choose fuel with higher amounts of ethanol per volume. However, meeting the 2017 and 2018 mandates of 15 billion gallons is more difficult, and the feasibility and the costs of expanding biofuels volumes is an important and current policy issue (see e.g., Pouliot and Babcock 2016 or Knittel et al. 2017). With the forecast for US gasoline consumption for 2018 of 143.3 billion gallons (US EIA 2018a), ethanol consumption through E10 will fall short of the mandate by 667 million gallons.

The approved alternative blends to increase ethanol consumption beyond E10 levels are E15 and E85,[2] which on average contain, respectively, 15 percent and 74 percent ethanol by volume. EPA is relying on increased E85 consumption to meet its ethanol blending mandate because so few stations are equipped to sell E15. The cost and feasibility of meeting the RFS blending targets depends on the number of motorists who have vehicles capable of refueling with E85 and who have access to an E85 station, and on their willingness to switch from E10 to E85. Fewer than 10 percent of US vehicles are flex-fuel vehicles (FFVs) that can use E85.

Recent studies of willingness to pay (WTP) for E85 have used either contingent valuations or discrete choice experiments. Jensen et al. (2010) find that consumers are willing to pay a premium for E85 made from corn ethanol of more than 13 cents per gallon compared to E10.[3] Petrolia et al. (2010) also find that consumers are willing to pay a premium for E85. The survey data show that motorists are

---

[1] It is beyond the scope of this article to describe the RFS program in all its details. For a good, succinct overview of the RFS, see Bracmort (2018).
[2] Blender pumps can deliver any blend between E10 and E85 by adjusting the ratio of the two.
[3] The authors also look into willingness to pay for E85 made using ethanol from switchgrass or wood residues. Ethanol made from these feedstocks is available only in very small quantities.

willing to pay, on a cost per mile basis, 12-15 cents more per gallon for E85 compared to conventional gasoline (E0). Estimation results in Petrolia et al. (2010) imply that motorists are willing to pay a premium between 3 and 6 cents per gallon for E85 over E10. Aguilar et al. (2015) use a discrete-choice experiment from a nationwide survey to estimate motorists' preferences for E0, E20, and E85. According to Aguilar et al. (2015), if these fuels were priced at parity on a cost per mile basis, they would respectively capture 21 percent, 17 percent, and 62 percent of the market by volume, signaling that the average motorist is willing to pay a premium for E85. These studies are likely subject to a hypothetical bias as respondents were not making actual fuel choices and several did not drive a vehicle that can run with a high-ethanol blend. The only study that uses revealed preference data to estimate demand for E85 in the United States is Anderson (2012) who uses data from a survey of fuel stations in Minnesota. However, Anderson (2012) cannot recover the full distribution of WTP for E85 because the energy-adjusted price of E85 was always greater than the price of E10.

To gain a better understanding of US motorists' preferences for E85, we followed the example of Salvo and Huse (2013), who conducted an intercept survey of motorists in Brazil. We conducted an intercept survey of US motorists with FFVs (flex motorists) at fuel stations in multiple states (Arkansas, California, Colorado, Iowa, and Oklahoma). We designed our survey to overcome circumstances not faced by Salvo and Huse (2013). Very few US stations sell E85, so many US flex motorists must incur additional costs (i.e., driving out of their way or forgoing a more convenient E10-only station) to refuel with E85. This implies that our sample of flex motorists have, on average, a higher WTP for E85 than the population of flex motorists. Survey questions allow us to correct for this selection bias.

We observed little variation in prices within some states during our survey, making identification of the complete distribution of preferences difficult based only on the motorists' fuel choices. We overcome this problem by combining revealed preference data with stated preference data. We observed motorists' actual fuel choices (from afar), and then we approached and presented them with a hypothetical set of prices and asked them to make a new (hypothetical) fuel choice. By combining the stated preference data with the revealed preference data, we can more precisely estimate the distribution of preferences for E85. However, because we did not randomize the hypothetical prices, but rather generated them to induce switching, we must carefully correct for the endogeneity problem created by our data collection method.

We find that the average flex motorist prefers E10 when prices for the two fuels are at parity on a cost-per-mile basis. However, preferences are diverse, fuel-switching behavior spans a wide range of relative prices, and about 25 percent of motorists choose E85 even when it is more expensive on a cost-

3

per-mile basis. The policy implication of our results is that the price of Renewable Identification Number (RIN) needed to induce enough consumption of ethanol to meet proposed blending mandates is higher than estimated in previous studies calibrated using preferences for E85 stronger than what we find (e.g., Pouliot and Babcock 2014).

Intercept Survey Design

We designed an intercept survey of flex motorists at fuel stations that sell E85 to obtain data on a broad range of factors that might affect WTP for E85 as a substitute for E10. The survey shares similarities with the survey of motorists at fuel stations in Brazil conducted by Salvo and Huse (2013). After first observing motorists' fuel choices from afar, we briefly interviewed the motorists. In almost every case, we did not detain the motorists longer than the time it took them to refuel. A supplementary appendix online contains the complete survey questionnaire.

Surveyed Fuel Stations

We obtained the cooperation of two E85 retailers to conduct our survey. We solicited the collaboration of other retailers but did not receive positive responses from them. The E85 stations we visited in Arkansas, Colorado, Iowa, and Oklahoma were operated by a retailer we will call 'Retailer Y', and the E85 stations we visited in California were operated by a retailer we will call 'Retailer Z'.

The installation of E85 pumps is costly (see e.g., Corts 2010 or Pouliot and Babcock 2014). For our study, we are concerned that the fuel stations surveyed are located in areas where motorists have preferences for E85 that are not representative of the population of motorists or that prices at a fuel station may be endogenous to the characteristics of the local population. We do not believe either to be a significant problem, especially at Retailer Y.

Retailer Y is a large retail fuel chain that sells motor fuel mostly in the Midwest. The retailer was an early mover in selling E85 and was still seeking to increase its offering at the time we conducted our survey. There was little to no competition for E85 near Retailer Y's fuel stations where we conducted the survey, but there was always nearby competition for E10. At the time of the survey, only two stations, which we label *Colorado Springs 2* and *Little Rock 3*, faced E85 competition within one mile. The retailer installed E85 pumps in new stations, at stations that needed renovations, provided these stations were of sufficient sizes, favoring states that offered incentives for E85 sales. Thus, local flex motorists' preferences for E85 was not a preponderant motive in the decision to install E85 pumps at a particular station but rather supply-side factors dominated. Retailer Y actively competed in the E10 market and its

prices for E10 were set to match nearby competition. Retailer Y set their E85 prices based on its cost of procurement of E85 at a given location plus a markup similar to the one for E10.

The decision process to install E85 pumps was different for Retailer Z. Retailer Z installed standalone E85 and biodiesel pumps alongside existing fuel stations. This required cooperating fuel stations to have available physical space, to agree to the installation of new pumps, and to obtain the necessary permits. Demand-side factors also mattered because Retailer Z sought to install pumps where there were enough potential consumers, and hence the pumps were located in large cities and in high-trafficked areas. At the stations we surveyed, there was nearby competition for E10 but no competition for E85 within one mile. The E85 price was set so that it was about 80 percent of the price of E10, in nominal value, as long as it covered cost. Retailer Z did not adjust the price of E85 with respect to local motorists' preferences but rather calibrated the price of E85 with respect to the price of E10. Retailer Z did not have any control on the price of E10 at stations where it installed pumps.

Our main concern in selecting when and where to conduct the survey was to obtain a sufficient amount of observations to justify the time we spent at a station. Hence, the stations surveyed were large, and we spent more time surveying motorists during commute hours. In California, we excluded certain fuel stations out of safety concerns. We conducted the survey in Iowa over the course of a few months. We surveyed fuel stations outside Iowa spending one week in Colorado, one week in Oklahoma, and one week in Arkansas. We spent one week in Northern California and one week in Southern California. We aimed to visit at least two or three stations in each city we visited with no target on the number of observations collected at a station. We spent a minimum of a day at each location.[4]

Intercept Survey Method

For each station we visited, we recorded station-level data including the station name and brand, the station address, prices of E10 fuels (usually regular, midgrade, and premium), and the price of E85. We made an effort to interview all flex motorists who pulled alongside any of the station's pumps. When a second flex motorist pulled up to a pump during an interview, we did not interview the second flex motorist. Instead, we completed the first interview and then waited to interview the next flex motorist who arrived. This sequencing rule avoided a possible selection bias by the interviewer. In practice, because the FFV-share of the vehicle fleet is small and the survey was over quickly, we managed to capture virtually all of the flex motorists who visited the E85 stations.

---

[4] At one fuel station in Colorado and one fuel station in Oklahoma, we spent only one day because we were able to collect at these stations only 11 and 12 observations respectively for an entire day of work.

We visually identified FFVs in two ways. First, many newer FFVs have a badge on the back (or in rare cases on the side) of the vehicle that indicates they are FFVs. Second, most FFVs have a yellow gas cap, a yellow ring, or a yellow sticker inside the gas door indicating that it is capable of using E85. In practice, identifying FFVs required the interviewer to walk around the pumps and closely inspect vehicles as they were refueling. A third way to tell whether a vehicle was an FFV was if the motorist chose E85. However, a few motorists made a refueling mistake by choosing E85 for a conventional vehicle not equipped to use it or had a vehicle with aftermarket modifications to use E85.[5]

Before talking to a motorist, the interviewer passively observed each motorist's fuel choice and vehicle characteristics, including vehicle make, model, vehicle type (car, truck, SUV, or van), the state on the license plate, whether the vehicle had an FFV badge, whether the vehicle had a yellow gas cap, and the gender of the motorist. The interviewer also recorded the transaction volume and expenditure after the motorist finished refueling.

The survey focuses on flex motorists. The implication is that the flex motorists we surveyed typically drove large and less efficient vehicles as we discuss later. We do not have priors regarding whether the perception of these motorists of biofuels is different from the population of all motorists. We discuss the possible selection problem from motorists' vehicle later in the text.

Survey Questions

Once a flex motorist made a fuel choice and began refueling, the interviewer approached and asked the motorist to participate in a short survey. The interviewer followed with questions about the motorist, motorist's awareness of E85, and motorist's opinions on topics that might explain the motorist's fuel choice.[6] A supplementary appendix online contains data and details regarding these questions.

---

[5] Over the course of conducting the survey, we learned that a small share of motorists have aftermarket modifications to conventional vehicles (not originally manufactured as FFVs) to use E85 because its higher-octane content can improve the vehicles' (racing) performance. In most cases, the vehicles are modified so that they can use either E85 or E10, but in rare cases, the vehicles are configured so that they can only use E85, and switching back to E10 requires modifying the vehicle.

[6] In one question, we asked about ownership of the vehicle. We are particularly interested in identifying government vehicles because federal government employees and government employees in certain states driving an FFV are required to refuel with E85 if available. Corts (2010) shows that government fleet adoption of FFVs, as mandated by the Energy Policy Act, increased the number of retail E85 stations. However, Corts (2010) cannot say whether the increase in E85 stations led to an increase in motorists purchasing FFVs. Corts (2010) notes that most FFVs in the dataset were purchased prior to E85 becoming available and that flex motorists perhaps did not know of their vehicles' capabilities.

We wanted to know whether the motorists we surveyed were random draws from the general population of flex motorists or if they were in our sample because they drove out of their way to the station specifically because it offered E85. To motorists who chose E85, we asked, "Did you choose to fuel at this station because it offers E85?" If they responded positively, we followed by asking, "How far out of your way did you have to drive?" We use responses to these questions to inform whether motorists who chose E85 self-selected into the sample. The next section shows the sample-selection problem and discusses how we use responses to these two questions to select estimation samples.

We asked questions to obtain stated preference (SP) data to complement the revealed preference (RP) data by proposing a single hypothetical price scenario to each motorist. For motorists who refueled with E10, the scenario was that we either increased the price of E10 or decreased the price of E85. Conversely for motorists who refueled with E85. The amount of the hypothetical price change was $0.25, $0.50, or $0.75 per gallon. Six versions of the questionnaire specified the direction of the price change and the amount. A supplementary appendix online shows the questionnaire for the case where we increased the price of the fuel chosen by $0.25 per gallon. When arriving at the fuel station, the interviewer would begin with one version of the survey selected at random and then proceeded through the sequence of surveys, thus randomizing the amount and the form of the price change for any particular interview.

Fuel Choice Model and Sample Selection

In this section, we describe motorists' fuel choices in a random utility model where motorists make their fuel decisions based on the ratio of the price of E85 to E10. In a supplementary appendix online, we offer an alternative model where motorists make their fuel decisions based on the difference in the price of E85 and E10. Self-selection manifests in our sample in two ways. A first self-selection problem stems from us interviewing only motorists who drive FFVs and these motorists may have stronger preferences for E85 than the average US motorist. The second self-selection problem stems from interviewing motorists at fuel stations that offer E85. We argue below that the first selection problem is likely small because supply-side incentives motivate sales of FFVs. We follow by describing how motorists who drive to a fuel station specifically for E85 self-selected into our survey. We then discuss how we obtain estimation samples to deal with the second selection problem.

A Model of Motorists' Fuel Choice

The random utility model below shows how motorists make their fuel choices based on the ratio of the energy-adjusted E85 price to the E10 price and motivates our empirical approach. The model assumes that the demand for fuel is perfectly inelastic in the short run. That is, motorists choose either E85 or E10 based on the E85 price ratio but the volume of fuel is price-independent. For now, we consider only motorists who do not incur a cost to access the E85 station. That is, for now we do not consider motorists' fuel station choice, which is at the heart of the selection problem. We use a subscript e to denote E85 and a subscript g to denote E10.

The indirect utility of flex motorist i derives from fuel j is

$$V_{ij}(p_j, \mathbf{x_i}, e_{ij}) = p_{ij}{}^{a_j} \cdot \mathbf{x_i}^{\mathbf{b_j}} \cdot \exp(e_{ij}), \qquad (1)$$

where e_{ij} is a type 1 extreme value random variable. The vector $\mathbf{x_i}$ includes motorist characteristics (e.g., age, gender, opinions) and possibly characteristics of a fuel station that may influence motorists' choices (e.g., signage, ownership). Taking logs on both sides, letting $a_e = a_g \equiv a$ so that prices affect indirect utility in the same manner as they are expressed in energy-equivalent terms, and letting $\mathbf{b} \equiv \mathbf{b_e} - \mathbf{b_g}$, we can write that a motorist chooses E85 if $\log V_{ie}(\cdot) \geq \log V_{ig}(\cdot)$ or

$$a \log(r_i) + \log(\mathbf{x_i}')\mathbf{b} \geq e_i,$$

where $r_i \equiv p_{ie}/p_{ig}$ is the E85 ratio and $e_i \equiv e_{ig} - e_{ie}$. Thus, preferences follow a logistic distribution, and the variables enter the model in logs.

Self-selection from vehicle choice

FFVs are alternate versions of conventional models, and the operation of an FFV is identical to the conventional version except that fuel efficiency with E85 is about 22 percent lower with E10.[7] For a particular vehicle make and model, consumers are not usually able to choose between flex and conventional versions. The incentives to manufacturers for selling FFV vehicles largely stem from a substantial credit from the US Corporate Average Fuel Economy (CAFE) standards for producing FFVs (see e.g., Anderson and Sallee 2011) rather than from consumer demand. The credit started declining in 2015 and will cease in 2020. The rule treats FFVs, up to an annual limit, as though they operate partially on E85, but the calculation of fuel economy uses the total miles the vehicle can travel per gallon of gasoline input (the ethanol fuel input is excluded in the fuel economy calculation). Anderson and Sallee

---

[7] The 22 percent fuel efficiency difference is an approximation based on the energy yields of ethanol and gasoline. For a given blending rate, the effective difference in fuel efficiency of E10 and E85 depends on car engines' ability to exploit the higher-octane level of E85 compared to E10. Nonetheless, the difference in fuel efficiency is near 22 percent across the vehicles currently on the market.

(2011) observe that "the AMFA formula's flexible-fuel credit mechanically treats inefficient models more generously than efficient ones, so automakers should prefer inefficient models, all else equal".

An early study by Rubin and Leiby (2000) reports that at least one-half of the penetration of alternative fuel vehicles, which include FFV, was due to a favorable regulatory treatment. Since then, the Energy Independence and Security Act of 2007 gave vehicle manufacturers increased incentives to produce FFVs. We are not aware of any recent studies that show the value of flex fuel capability to motorists and thus we cannot exclude that certain motorists purchased FFVs to be able to use E85.

Three arguments support that the CAFE incentive explains most of the ownership of FFVs. The first argument is the composition of vehicles offered as FFVs across manufacturers. Car manufacturers who mostly produce fuel-efficient vehicles such as Honda, Kia, Subaru, and several others do not offer FFVs in the United States. The only FFVs that Toyota and Nissan offer are their largest trucks (Tundra and Titan, respectively) and largest SUVs (Sequoia and Armada, respectively). American car manufacturers such as Ford, GM and Chrysler offer FFVs for their larger vehicles (see Alternative Fuels Data Center 2017). The second argument is the rise and the decline in the number of FFVs that car manufacturers have made available in the United States. Data from US EIA (2018b) show that when the 2007 standards became effective in 2008, the number of new FFVs made available was 1.18 million. That number declined the following year because of the recession but then constantly increased up to 2.66 million by 2013. As the CAFE credits expired in 2014, the number of FFVs made available began to decline with 1.44 new FFVs made available in 2016, about the same number as in 2010. The third argument is that motorists have been purchasing FFVs even though most fuel stations did not offer E85. Pouliot and Babcock (2014) show that only a small fraction of flex motorists had access to E85 at a nearby fuel station. Hence, the possibility of purchasing E85, or other high blends, could not be an important factor in the decision of a motorist to purchase an FFV. These arguments support that the CAFE credits are the main reasons for the growth in the number of FFVs in the United States and explain why flex fuel capability has been offered for less fuel-efficient models. Therefore, the consumers who purchase FFVs are generally those who demand pickup, SUVs, and other larger light-duty vehicles.

Self-selection from fuel station choice

For mail, phone, or internet surveys, it is possible to develop rules to obtain a sample that is representative of the population. In contrast, for a survey at fuel stations like ours, we only observe flex motorists who choose to refuel at that station. For such on-site surveys, if showing up at the survey location correlates with attitudes toward the topic of the survey, then there will be a sample selection

problem. In our case, offering of E85 may motivate patronage to a station, meaning that certain motorists may self-select into our sample.

There are reasons to believe that the sample selection problem could be important. In the United States in 2014 and 2015 (when we conducted our survey) there were about 2,700 fuel stations that offered E85 while E10 was available at all of the nearly 110,000 fuel stations (Pouliot and Babcock 2014; Clinton et al. 2017). While motorists could access E10 at any station along their normal driving routes, most could not access E85 at their preferred or most convenient station, thus many had to deviate from normal driving routes to access E85. Thus, many flex motorists we observed refueling with E85 likely incurred costs associated with forgoing E10-only stations.

To formalize the description of the self-selection problem, we can extend the choice model above to consider that motorists simultaneously choose a fuel station and a gasoline blend. We identify the E85 station where the survey is conducted by k and other stations by $-k$. The indirect utility of motorist i for fuel j at station k is

$$V_{ij}\big(p_{jk} + c_{ik}, \mathbf{x_i}, e_{ij}\big), \tag{2}$$

where c_{ik} is the cost to motorist i of accessing a station k, expressed in dollars per gallon. The difference between equation (2) and equation (1) is that the price is now station specific and accessing a particular fuel station may be costly for some motorists because of driving costs (e.g., Houde 2012, Wolff 2014). In (2), motorists augment the fuel price with the cost to them of accessing the fuel.

E10 is available at every fuel station, so motorists can always access E10 at their most convenient or preferred station. The cost of accessing station k does not necessarily depend only on a motorist's home location but also on the motorist's regular driving habits. For instance, station k can be miles away from a motorist's home, but if station k is in a convenient location along a regular commute, the cost of accessing the station is low. Without loss of generality, we can normalize the cost for motorists to access their most convenient or preferred station to zero, i.e., $c_{ik} = 0$.

We assume, to simplify the discussion, a motorist who cares only about the cost per mile and considers purchasing fuel at station k and $-k$. Consequently, we adjust the price of E85 to be energy equivalent to E10. We refer to the net price as the price of a fuel plus the access cost. Station k is the most preferred station for this motorist i such that $c_{ik} = 0$, everything else being equal. The motorist will purchase E85 at station k if the following conditions hold: 1) $p_{ek} \leq p_{gk}$; 2) $p_{ek} \leq p_{g-k} + c_{i-k}$; and 3) $p_{ek} \leq p_{e-k} + c_{i-k}$. There is no selection problem in this case because the motorist purchase of E85 at her preferred station.

Alternatively, if motorist i does not have access to E85 at her most convenient fuel station k such that $c_{ik} > 0$, she will refuel with E85 at fuel station $-k$ if two conditions hold: 1) $p_{e-k} \leq p_{g-k}$; and 2) $p_{e-k} + c_{i-k} \leq p_{gk}$. Here, the motorist purchases E85 at station $-k$ even though she must incur an additional cost c_{i-k} to access E85 at that station. In this case, the decision of the motorist to refuel at station $-k$ correlates with its offering E85.

The discussion above held everything but prices constant and did not consider the preference of the motorist. Motorists' preferences for E85 increase the sample selection problem. Figure 1 illustrates from a simulation how the self-selection manifests in data collected at a fuel station. The simulation uses a utility specification consistent with equations (1) and (2).[8] Figure 1 shows histograms of preferences for motorists who visit an hypothetical fuel station. The vertical black line is the (log) net price ratio and serves as the cut-off point where motorists on the left purchase E10 and motorists on the right purchase E85. For the two categories of motorists, the shaded areas sum to the number of motorists. In green are the flex motorists who visit the fuel station at zero cost. In purple are the motorists who drive to the fuel station specifically because it offers E85 such the distribution of their preferences strictly lies on the right of the vertical black line. These motorists have strong preferences for E85 and forgo their most conveniently located fuel stations to access E85. They represent the new business that the fuel station acquires because it offers E85 and are the motorists who self-select into our survey.

Estimation Samples

This section explains how we select observations to create estimation samples that approach a random sample from the population.

Recall that we asked E85 motorists two questions to inform self-selection. The first question was, "Did you choose to fuel at this station because it offers E85?" To the motorists who answered "yes," we followed by asking, "How far out of your way did you have to drive?"[9] We use responses to these questions to select estimation samples. The first sample includes all the E10 motorists and only the E85 motorists who answered "no" to the first question. Using motorists' responses to the first question to select an estimation sample may be too restrictive. Indeed, it is possible that some motorists

---

[8] The code for the simulation is available in a supplementary appendix online.
[9] We designed the survey questions to be simple and short. In hindsight, a better version of the question would have been, "Imagine that all stations offer E85 and E10. Would you still be buying E85 at this station or another station closer to your home, work, or usual driving routes?"

who answered "yes" to the question about whether their patronage was motivated by the offering of E85 would have chosen the same station even if every station offered E85. Many motorists answered the follow-up question about how far out of their way they drove to access E85 with, "Not at all," "I didn't," or "Zero." A better estimation sample therefore likely consists of all of the motorists who chose E10, all of the motorists who chose E85 and answered "no" to the first question, and motorists who chose E85 and answered "yes" to the first question and also answered (some form of) "zero" to the second question. Our second estimation sample includes all of these motorists.

The objective in selecting these two estimation samples is to capture only the E85 motorists who did not self-select into the sample. We observe two groups of motorists refueling at an E85 station: 1) motorists for whom the station is the most preferred station independent of the station offering E85 and 2) motorists for whom some other E10-only station is the most preferred station, but the motorists forgo that station because they want E85. We want our estimation sample to capture only motorists in the first group, i.e., the motorists who are part of the green distribution in Figure 1. We employ two estimation samples because we cannot determine with certainty which of the motorists should be part of our estimation sample. we compare models where the estimation sample includes the E10 motorists and only the E85 motorists who answered "no" to the first question to models where the sample also includes E85 motorists who answered "yes" to the first question and "zero" to the second question.

The estimation sample that best approaches a random sample of the population of flex motorists likely lies between our two samples. The first sample is possibly too restrictive because it does not include enough E85 motorists and thus will cause us to underestimate WTP for E85. The second sample might be too inclusive and will cause us to overestimate WTP for E85. Thus, by using the two estimation samples we will estimate bounds on the population parameters and WTP distribution.

For both samples, we assume that E85 at a station did not motivate the E10 motorists' patronage to that station. It is possible that a motorist drove to a fuel station for E85 but, after observing prices, chose to refuel with E10. We believe this happens rarely and that self-selection by E10 motorists is much less of a problem than self-selection by E85 motorists. In talking to the E10 motorists about hypothetical prices and choices, refueling habits, and knowledge of E85, we never encountered motorists who claimed this behavior and motorists were generally aware of the prevailing fuel prices and seemed to make their fuel choices before arriving at the station. Therefore, we treat E10 motorists as though they did not self-select into our samples. If a few E10 motorists self-selected into our sample of motorists, this will cause a downward bias in the estimation of WTP for E85.

One could consider modeling the selection problem so that the empirical model uses the full sample of data. However, this would require either knowing how much it costs motorists to drive to an out-of-the-way fuel station or making strong assumptions to make identification of transportation costs possible. One important difficulty is that in the data the distance driven to access E85 correlates positively with the motorists' choice of E85. In the population for FFV motorists, this correlation is negative. Instead of relying on strong assumptions to identify transportation costs, we elect to work with smaller data samples that are more closely representative of the population of flex motorists.

If driving costs were zero, the distribution of preferences among motorists who self-selected into the sample would be the same as the distribution of preferences of motorists randomly drawn from the population who chose E85. That is, in Figure 1, the shape of the histogram in purple would be the same as the histogram in green on the right of the black vertical line. In such a world, we could include the observations where motorists self-selected into our sample and a correction for choice-based sampling. The literature offers estimators for this stratified sampling problem (see e.g., Manski and Lerman 1977, Manski and McFadden 1981, Imbens 1992, Imbens and Lancaster 1996). However, driving costs do not equal zero and can be quite significant (e.g., Houde 2012, Wolff 2014). The difference in the shape of the distribution of preferences for the two types of motorists refueling with E85 violates a fundamental assumption in models of choice-based sampling.

Description of the Data

We collected 972 observations of flex motorists from 17 E85 stations in six urban areas between October 2014 and April 2015.[10] In chronological order, the urban areas we visited were Ames/Des Moines, Iowa; Colorado Springs, Colorado; Tulsa, Oklahoma; Little Rock, Arkansas; Sacramento, California; and Los Angeles, California. We personally collected most of the observations. A small team of undergraduate students helped collect observations in the Ames/Des Moines area. In each urban area, we visited between two and four stations and collected around 100 or more observations.

From the initial 972 observations of motorists refueling their FFVs, we remove 79 observations where motorists chose not to or were unable to complete or participate in the survey. This represents a total non-response rate of 8 percent. In addition, we removed 12 observations for which we do not have

---

[10] We collected 994 observations but 22 were from conventional vehicles with aftermarket modifications to use E85. We observed these vehicles refueling with E85, but we exclude them because we can only identify these vehicles as flex when the motorists choose E85.

SP data because we asked the question incorrectly or the motorist was unable to answer. That leaves us with an initial sample of 881 complete observations before we address the sample-selection problem.

Price and Fuel Choice Data

Table 1 summarizes the fuel choice data broken down by station, urban area, and retailer. We adjust the E85 price so that it is energy-equivalent with the E10 price by multiplying it by 1.28=1/0.78. With this normalization, a dollar spent on E85 yields the same mileage as a dollar spent on E10.[11] In the entire sample of 881 flex motorists, the average E85 price was $2.81 per gallon, and the average E10 price was $2.58 per gallon. The average price ratio (defined as the E85 price divided by the E10 price) was 1.09, the minimum price ratio was 0.94 in Des Moines and the maximum price ratio was 1.28 in Colorado Springs. We observed greater price variations in Des Moines because we spent more time there conducting surveys. 431 (49 percent) flex motorists chose E85 while 450 (51 percent) chose E10.[12]

On average, E85 fuel prices were more favorable toward E85 at Retailer Z in California with an average E85 ratio of 1.06 compared to an average E85 ratio of 1.10 at Retailer Y. However, we observed more variation in the price ratio at Retailer Y where it ranged between 0.94 and 1.28 compared to a smaller range between 1.01 and 1.10 at Retailer Z. We observed that 89 percent of motorists chose E85 at Retailer Z's locations, and 35 percent of motorists chose E85 at Retailer Y's locations. Retailer Z's pumps served a larger share of the local flex motorists because E85 stations were less common in California. In addition, Retailer Z had previously run promotions providing special fuel cards and other incentives to local flex motorists, marketing E85 as a clean-burning, high-performance fuel. We do not have evidence that any particular promotion was taking place while we conducted the survey.

For Retailer Y, E85 prices were lowest in Iowa, where the average E85 ratio was 1.07, nearly the same as the average price ratio observed at Retailer Z. The share of flex motorists who chose E85 in Iowa was 42 percent, less than half of what we observed at E85 stations in California. We suspect that one reason for the difference is that stations that offer E85 are more common in Retailer Y's areas. Thus, local flex motorists with high WTP for E85 can possibly choose between multiple E85 stations during their regular commute. Price ratios within states are comparable across stations but their shares of

---

[11] This is true assuming that E85 contains 74 percent ethanol and 26 percent gasoline as assumed by US EIA (2017). The actual ethanol concentration in E85 varies slightly across states and over time.
[12] Among the 450 flex motorists who chose E10, 414 (92 percent) chose regular grade 87 octane (85 octane in CO), 24 (5 percent) chose midgrade, and 12 (3 percent) chose premium.

motorists who purchase E85 differ. This might be due to station-specific idiosyncrasies or randomness in the motorists interviewed given that the number of observations per station is not very high.

The column "% of E85 who did not drive" in Table 1 shows the share of E85 motorists who responded that they did not drive to the fuel station because it offers E85 and informs our first and more restrictive estimation sample. Using this first rule leaves us with a sample of 479 observations with 29 motorists choosing E85. At some stations, all E85 motorists reported that they had driven to the station specifically for E85. The second rule is more inclusive and uses observations where motorists answered zero to the question about the distance they drove to fuel with E85. The column "% of E85 who did not drive" is the share of E85 motorists who responded "no" to the first question and responded "zero" to the second and informs our second estimation sample. Only at one station were there no E85 motorists who qualified for the second estimation sample. This second rule gives us a sample of 670 observations with 219 motorists choosing E85.

Figure 2 shows the share of flex motorists by state who chose to refuel with E85 (RP choice) as a function of the price ratio for the with 670 observations. The corresponding figure with the sample of 479 observations is similar and not shown. The regression line shows that the share of motorists who chose E85 declines with respect to the price ratio, as expected. Observe that motorists in California appear to have stronger preferences for E85. The share of motorists who refuel with E85 in California is above 50 percent for all observed values of the price ratio. In Iowa, the share of motorists who purchased E85 was smaller than 50 percent even when the price ratio was relatively low.

Figure 2 also shows that for Arkansas, Colorado and Oklahoma, there is little variation in the price ratio in the RP data compare to Iowa and California where we spent more time. Although there is little variation in the price ratio in certain states, we are still able to estimate state-specific fixed effects because there is heterogeneity in motorists' choices, characteristics, and opinions. Estimation of the coefficients for the price ratio relies on variations in the price ratio within states.

Figure 3 shows in two panels the share of motorists who chose to refuel with E85 as a function of the hypothetical price ratio for the sample with 670 observations. The hypothetical price scenario that we presented to motorists was conditional on their fuel choice, and accordingly, we present the shares conditional on motorists' fuel choice (RP choice). Observe that in both panels, variations in hypothetical prices are much greater than the variation in prices in Figure 2. This is true both within and across states, which will allows us to estimate more precisely how motorists' choices respond to prices. Panel a) of Figure 3 is for observations we collected at Retailer Y and shows that the share of motorists who chose E85 declines with respect to the hypothetical price ratio regardless of whether motorists' RP

choice was E10 or E85. Panel b) is for observations we collected at Retailer Z. The share of motorists who chose E85 declines with respect to the hypothetical price ratio for those who refueled with E85 but unexpectedly increased for those who refueled with E10.

Motorists' Knowledge

In this section, we discuss motorists' knowledge of E85 from certain questions in our survey. A supplementary appendix online offers a more comprehensive summary.

A basic knowledge issue that might limit expansion of E85 consumption is flex motorists knowing that they can use E85 and that it is available at the some fuel stations. We asked three questions to E10 motorists to measure their knowledge of the flex-fuel capability of their vehicle and the existence of E85. Table 2 shows a summary of the E10 motorists' responses. Out of the 450 E10 motorists, 65 (14 percent) did not know that they drove a flex vehicle. We only asked the following two questions to motorists who knew they were driving an FFV because we assumed that motorists who did not know they drove an FFV had never refueled with E85 and were not considering refueling with E85. Only 38 percent of the E10 motorists who knew they were driving an FFV had previously refueled with E85, and 73 percent of them knew the fuel station offered E85. These responses show a certain level of ignorance among flex motorists about the characteristics of their vehicles and the availability of E85.

We were also interested in comparing the attitudes of motorists toward E10 and E85 according to their fuel choice and across states. Table 3 summarizes motorists' responses to five opinion questions. We asked motorists to respond with either "Ethanol", "Gasoline", or "No difference", but we also accepted "Don't Know" as an answer. Although the questions specifically asked about ethanol and gasoline, for the purpose of this discussion we will refer to the fuels as E85 and E10 as opinions about ethanol and gasoline likely extend to E85 and E10.

E85 motorists answered in a greater proportion compared to E10 motorists that E85 is better for the environment, the engine, the economy, for national security and that it yields more mile per gallon. This is true in the aggregate but also in almost every case for individual states. This indicates, not surprisingly, that E85 motorists have an attitude more favorable toward E85 compared to E10 motorists.

In the three questions about which fuel is better for the environment, the engine, and the economy, the differences in opinions across states are especially apparent. In general, a greater share of flex motorists we surveyed in Iowa and California believe that E85 is better for the environment, for a vehicle's engine, and for the economy, while motorists in Arkansas and Oklahoma have a much less favorable opinion of E85, and motorists in Colorado are somewhere in between. For some flex motorists

we surveyed, the question about national security elicited more confusion rather than an actual response. In 2006, national security and independence from foreign oil were touted as reasons to support the biofuels mandates, but the cause seems to have lost its importance since. As with the other questions, motorists in Iowa and California favor E85 for national security more than motorists in the other states, but many motorists answered, "No difference" or, "Don't know".

The last question of the survey asked which fuel yields more miles per gallon. Responses to the question did not differ by much across states except for California where a smaller share of respondents indicated that E10 was the most efficient fuel.

Responses to questions about the environment, the economy and national security are subjective as there are no definitive answers to these questions. For example, economists have long been debating whether the RFS indeed reduces greenhouse gas emissions. Responses to these three questions reveal the motorists attitude toward E10 and E85. In contrast, questions about which fuel is better for the engine and which fuel yields more mile per gallon are factual and reveal ignorance by certain motorists. For an FFV, there should be no difference between E10 and E85 on engine-wear, and E10 gives more miles per gallon because it contains more energy than E85. Responses to these two questions indicate that a significant share of flex motorists have knowledge gaps about the basic operation of their vehicle and the fuels it can use.

Ignorance about the energy difference of the two fuels likely explains why some motorists drive miles out of their way or wait in line to refuel with E85, which we observed on occasions, especially in California. We also asked motorists in a follow-up question to approximate the perceived energy difference between the two fuels. Some motorists responded with an accurate answer saying that E85 gets about 75-80 percent of the miles per gallon of E10. Some approximated higher energy for E85 in the 90 percent range and some approximated the energy ratio to be as low as 50 percent. Interestingly, many of the flex motorists who chose E85 demonstrated that they understood that E85 was more expensive on an energy-equivalent basis. Some chose E85 for reasons other than the price, while others simply did not bother to calculate the energy-equivalent fuel costs. Many flex motorists said something along the lines of, "I did the math once and figured that I need a $0.60 per gallon discount on E85 for it to be worth it," and now they make their fuel choice based on some rule-of-thumb or routine.

Empirical Models

In this section, we describe the three empirical models we use to obtain estimates of the distribution of WTP for E85. In the description of the empirical models, we use a generic notation where $\Lambda(\cdot)$ is the

cumulative logistic distribution, z_i is the vector of independent variables that includes the log of the E85 ratio, and θ is the vector of parameters to estimate. We write that y_i equals one for a motorist who refuels with E85, and y_i equals zero for a motorist who refuels with E10.

Maximum Likelihood

The first model we consider is the standard maximum likelihood estimator (MLE). Under standard assumptions, z_i is exogenous so we can consistently estimate θ by maximizing the conditional log-likelihood given by

$$logL = \sum_{i=1}^{N}\{(1 - y_i)\ln[1 - \Lambda(z_i'\theta)] + y_i\ln[\Lambda(z_i'\theta)]\}. \tag{3}$$

This is the log-likelihood function we use to estimate models that use only the RP data.

Augmenting RP Data with SP Data

When using our strictest sample-selection rule, we only have 29 observations where motorists choose E85 in the RP data. One solution to the rare-choice problem is to augment our RP data with SP data. As we will show, the MLE in (3) is not a correct approach for estimation of models that add SP data.[13] Moreover, as Figures 2 and 3 show, the addition of the SP data gives us more variation in relative prices for E10 and E85, which allows us to more precisely estimate motorists' responses to prices.

SP data have been used to complement RP data in previous studies to increase the number of observations and expand the choice set to include alternative(s) not otherwise available. The traditional method to combine RP and SP data in the transportation and the environmental economic literature has been described by Ben-Akiva and Morikawa (1990), Hensher and Bradley (1993), Adamowicz et al. (1994), and Hensher et al. (1999). The intuition behind the traditional approach is that the unobserved factors are different for the two types of data. To account for this, the RP and SP data are stacked together and the empirical model allows for different intercept and scale parameters for the distributions of the error terms for the SP and RP data. The traditional approach is appropriate when the attributes of the hypothetical choices in the SP data collection are independent of the RP choices so that respondents' unobservable characteristics do not correlate with the hypothetical options.

Train and Wilson (2008) and Train and Wilson (2009) consider SP data constructed from RP choices. They refer to these data as "SP-off-RP" data. The distinction from the traditional SP data is that the hypothetical choice scenario depends on the consumer's observed choice, introducing endogeneity.

---

[13] We could use the finite-sample correction proposed by King and Zeng (2001) as a solution to the rare-choice problem. However, we do not find that the correction has a significant impact on our results.

Recall that if we observed a motorist choosing E10, we offered hypothetical prices that lowered the relative price of E85. Conversely, to a motorist choosing E85, we offered hypothetical prices that lowered the relative price of E10. A motorist's RP fuel choice depends on both observed characteristics and unobservable factors. The same unobservable factors that affect the motorist's observed RP fuel choice carry over to the SP experiment, so the unobserved factors in the SP experiment correlate with the hypothetical prices, and we need to account for this when incorporating the SP data.

We define the (log) utility that flex motorist i derives from fuel j in the SP experiment as

$$\log W_{ij}(\dot{\mathbf{z}}_{ij}, e_{ij}, \eta_{ij}) = \dot{\mathbf{z}}_{ij}'\boldsymbol{\theta}_j + e_{ij} + \eta_{ij}$$

where $\dot{\mathbf{z}}_{ij}$ includes the log of the hypothetical price of fuel j and η_{ij} is a generalized extreme random variable with scale $(1/\zeta)$ that captures additional unobservable aspects of the SP scenario not present in the RP scenario. In the words of Train and Wilson (2008), η_{ij} accounts for the quixotic aspects of the SP task not present in the RP scenario. Note that in the utility function for the SP data, the relationships between both the observable and unobservable factors that determine the utility in the RP data in $V_{ij}(\cdot)$ in equation (1) are preserved. This means that the unobservable e_{ij} term for motorist i that affects the RP choice carries forward to the SP choice. The total unobservable error term in the SP model is $e_{ij} + \eta_{ij}$, where e_{ij} derives from the motorist's RP choice. The hypothetical prices $\dot{p}_{ie}$ and $\dot{p}_{ig}$ are endogenous because they are correlated with the total error term $e_{ij} + \eta_{ij}$.

A motorist chooses E85 in the hypothetical price scenario if $\log W_{ie}(\cdot) \geq \log W_{ig}(\cdot)$, or

$$\zeta(\dot{\mathbf{z}}_i'\boldsymbol{\theta} - e_i) \geq \eta_i,$$

where $\dot{\mathbf{z}}_i$ includes the log of the hypothetical E85 ratio and $\eta_i \equiv \zeta(\eta_{ig} - \eta_{ie})$ is symmetric with a mean of zero and follows a logistic distribution. The ζ term can be estimated and scales the logistic distribution of η_i to have a scale of one. The probability that a motorist chooses E85 in the SP choice experiment is $\Pr(\text{SP E85}_i) = \Lambda(\zeta[\dot{\mathbf{z}}_i'\boldsymbol{\theta} - e_i])$, and the probability that a motorist chooses E10 in the SP experiment is $\Pr(\text{SP E10}_i) = 1 - \Lambda(\zeta[\dot{\mathbf{z}}_i'\boldsymbol{\theta} - e_i])$. The joint probability of a motorist's specific RP and SP choice combination is the product of the probabilities of the RP choice and the conditional probability of the SP choice (conditional on the RP choice). We write $\dot{y}_i = (\dot{y}_{i1}, \dot{y}_{i2})$ where $\dot{y}_{i1}$ is the RP choice and $\dot{y}_{i2}$ is the SP choice. The SP-off-RP likelihood function is

$$L = \prod_{\dot{y}_i=(0,0)} Pr(\text{RP E10}_i)Pr(\text{SP E10}_i|\text{RP E10}_i) \prod_{\dot{y}_i=(0,1)} Pr(\text{RP E10}_i)Pr(\text{SP 85}_i|\text{RP E10}_i) \quad (4)$$
$$\prod_{\dot{y}_i=(1,0)} Pr(\text{RP E85}_i)Pr(\text{SP E10}_i|\text{RP E85}_i) \prod_{\dot{y}_i=(1,1)} Pr(\text{RP E85}_i)Pr(\text{SP E85}_i|\text{RP E85}_i).$$

The e_i's that enter the SP probability expressions are not observed but we know their conditional distributions so we can integrate to calculate the expected value. For example, the logit probability of a motorist choosing E85 in the SP experiment conditional on that motorist choosing E85 in the RP data is

$$\Pr(\text{SP E85}_i | \text{RP E85}_i) = \Lambda(\zeta[\dot{\mathbf{z}}_\mathbf{i}'\boldsymbol{\theta} - e_i | e_i \leq \dot{\mathbf{z}}_\mathbf{i}'\boldsymbol{\theta}]),$$

which we can write as

$$\Pr(\text{SP E85}_i | \text{RP E85}_i) = \int \Lambda(\zeta[\dot{\mathbf{z}}_\mathbf{i}'\boldsymbol{\theta} - e_i]) \lambda(e_i | e_i \leq \dot{\mathbf{z}}_\mathbf{i}'\boldsymbol{\theta}) de_i,$$

where $\lambda(\cdot)$ is the marginal density of the logistic distribution. We evaluate the integrals by simulation, taking draws of e_i from its conditional density following the method described by Train and Wilson (2009). We calculate the probability $\Lambda(\cdot)$ for each draw and average the results. We estimate the parameters by maximizing the log of the likelihood function in equation (4) using 1,000 conditional logistic draws for each observation.

Endogenous Stratification

Lastly, we estimate a version of the SP-off-RP model that accounts for endogenous stratification. Because E85 contains about 78 percent of the energy of E10, motorists who repeatedly refuel with E85 must refuel nearly 30 percent more frequently than they would with E10. For our survey, the implication is that we interviewed more E85 motorists then we would have with a random sample from the population. A literature proposes estimators for stratified sampling (e.g., Manski and Lerman 1977, Manski and McFadden 1981, Imbens 1992; Imbens and Lancaster 1996). For our purpose, we can write the weighted likelihood function as

$$wL = \prod_{\dot{y}_i=(0,0)} Pr(\text{RP E10}_i)^{w_{i0}} Pr(\text{SP E10}_i | \text{RP E10}_i)^{w_{i0}}$$

$$\prod_{\dot{y}_i=(0,1)} Pr(\text{RP E10}_i)^{w_{i0}} Pr(\text{SP 85}_i | \text{RP E10}_i)^{w_{i0}} \qquad (5)$$

$$\prod_{\dot{y}_i=(1,0)} Pr(\text{RP E85}_i)^{w_{i1}} Pr(\text{SP E10}_i | \text{RP E85}_i)^{w_{i1}}$$

$$\prod_{\dot{y}_i=(1,1)} Pr(\text{RP E85}_i)^{w_{i1}} Pr(\text{SP E85}_i | \text{RP E85}_i)^{w_{i1}},$$

where w_{i0} is the weight applied to observations where motorists refuel with E10 and w_{i1} is the weight applied to observations where motorists refuel with E85. From the relative energy content of E10 and E85, we set $w_{i0} = 1$ and $w_{i1} = 0.78$. The intuition is that, when estimating the share of motorists who choose E85, we give less weight to observations of E85 motorists to compensate for them being overrepresented in our sample from having to visit the station more often.

An underlying assumption of the endogenous stratification correction is that motorists who refuel with E10 and E85 have similar behaviors when it comes to driving intensity and filling up their tanks. We did not collect information about whether motorists filled up their tanks or a justification for

the volumes they purchased. However, we recorded the volume purchased and total expenditure. We show in a supplementary appendix online histograms of the volumes of fuel purchased and fuel expenditures. There is no apparent difference in the behavior of E10 and E85 motorists, especially for the volumes purchased, hence supporting the use of weights for the relative energy content.

Estimation Results

We begin by discussing marginal effects to show which factors influence motorists' fuel choices. We then calculate the parameters of the logistic distribution of WTP implied by the regression results.

The regression models use explanatory variables for vehicle ownership (personal, government, company, other), vehicle type (car, truck, SUV, van), whether the vehicle had an FFV badge, number of miles driven per year, gender, age, opinions about which fuel is better for the environment, the engine, the economy, and national security, opinion about which fuel yields more miles per gallon, and the state where the station was located (Arkansas, California, Colorado, Iowa, or Oklahoma). We do not include variables that describe the characteristics of the fuel stations because the state dummies summarize most of that information. We show in a supplementary appendix online results for models where we use station-specific fixed effects to capture unobserved heterogeneity across stations. The results are very similar to those we describe in what follows.

Marginal Effects

We estimate six models from three empirical models and two estimation samples. The three empirical models are: A) MLE with the RP data only (RP data); B) SP-off-RP method (SP-off-RP); and C) SP-off-RP with correction for endogenous stratification (Endo strat). The estimation samples contain 479 and 670 observations depending on the rules we use to determine whether E85 motorists self-selected into the sample. We summarize the regression results by showing marginal effects in Table 4 for the sample with 479 observations and in Table 5 for the sample with 670 observations.

In all models, for both samples, motorists are more likely to buy E85 when its price drops relative to E10. For the sample with 479 observations, the marginal effects for the log price ratio are stronger for Model A that uses the RP data only and are comparable between Model B that uses the SP-off-RP data and Model C that uses endogenous stratification weights for the fuels' energy contents. The marginal effects for the log ratio are also largest in Model A for the sample with 670 observations and are smallest in Model C. In Model A, for both samples, the marginal effects for the price ratio are not statistically different across models largely due to the large standard errors. Incorporating SP data in

21

Models B and C meaningfully improves precision in measuring price responsiveness. The marginal effects for the price ratio are much larger in the models with 670 observations than in the models with 479 observations. The sample with 670 observations includes more E85 motorists, with some of them driving more likely to drive out of their way when the price of E85 is low.

The interpretation of the marginal effect for the log of the price ratio, calculated as $\partial Pr(E85)/\partial \log(r)$, is the change in the probability of purchasing E85 from a one percent increase in the price ratio. The marginal effect from Model C with the sample with 479 observations implies that the purchase probability of E85 declines by 0.00242 from a one percent increase in the price ratio. In Model C with 670 observations, a one percent increase in the price ratio causes a decline in the purchase probability of E85 by 0.00546. We can also interpret our results in terms elasticities, which we calculate as $\partial Pr(E85)/\partial \log(r) \cdot Pr(E85)$. In the sample with 479 observations, 29 motorists chose E85, giving us a baseline probability of 6.05 percent. From the marginal effects in Table 4, the elasticities for Models A, B, and C are respectively -5.65, -4.12 and -4.00. In the sample with 670 observations, 220 motorists chose E85, giving us a baseline probability of 32.84 percent. From the marginal effects in Table 5, the elasticities for Model A, B, and C are respectively -2.07, -1.88 and -1.66 for the sample with 670 observations. These results are consistent with expectations that the demand for E85 is elastic because flex motorists can readily switch from E85 to E10. Our elasticity estimates are similar to Aguilar et al. (2015) who find elasticities between -3.00 and -2.00 and Anderson (2012) who finds an own price of elasticity of demand for E85 of negative 3.20-3.80.

A few other variables are statistically significant. Notably, the fixed effect for motorists in California is positive, statistically significant and large. The only vehicle ownership effect that is significant is for motorists who drive a government vehicle; they are more likely to purchase E85 as expected. Motorists who believe that ethanol is better for the environment, for a car engine, for the economy, and who believe that there is no difference in miles per gallon between E85 and E10 are more likely to purchase E85. In the sample with 479 observations, motorists who believe that gasoline is better for the environment or that ethanol and gasoline have the same impacts on the environment are less likely to purchase E85. However, in the sample with 670 observations, these marginal effects tend to be not statistically significant.

For variables other than the price ratio, the precision gains of using SP data are small. Indeed, the SP experiment does not add information for factors other than the price, and therefore we did not expect that the SP data would meaningfully improve precision for variables other than the price ratio.

Willingness to Pay Distribution Parameters

It is useful to describe WTP for E85 versus E10 in terms of distribution parameters because it is then possible to use these parameters in calibrated models of ethanol demand such as in Pouliot and Babcock (2014). Below, we report the scale and shape parameters that summarize the log-logistic distribution of WTP. The scale parameter represents the median WTP and expresses the intensity of preference for E85 relative to E10, and the shape parameter describes how diverse preferences for E85 are.

We calculate the scale parameter of the log-logistic distribution of WTP as $\tilde{\rho} = exp\left(\overline{\log(\mathbf{x_i})'}\,\hat{\mathbf{b}}/-\hat{a}\right)$, where we use the bar to denote a mean. We calculate the shape parameter as $\tilde{s} = -\hat{a}$.[14] The scale parameter, $\tilde{\rho}$, is the median WTP expressed in terms of the price ratio. When $\tilde{s} > 1$, the log-logistic is unimodal and the expression for the mean is $\frac{\tilde{\rho}\pi/\tilde{s}}{\sin(\pi/\tilde{s})}$. We calculate the standard errors for the distribution parameters using the delta-method.[15]

We compare willingness to pay at Retailer Y and Retailer Z using the estimates of scale parameters. We find small differences in the scale parameters for the four states where we conducted interviews at Retailer Y as implied by the estimated fixed effects in Table 4 and 5 show. However, these differences are trivial when compared to the difference in the value of the location parameter for motorists in California. Thus, we will focus the discussion on summarizing the WTP distribution at all of Retailer Y's locations collectively relative to the WTP distribution at Retailer Z's locations in California.

Recall that the first estimation sample uses a stricter selection rule for E85 motorists and likely underrepresents the population of E85 motorists. In contrast, the second sample uses a looser selection rule and possibly over-represents the population of E85 motorists. The more E85 motorists an estimation sample contains, the higher the mean WTP we will find. Hence, we can think of our estimates from the more restrictive sample with 479 observations as a lower bound on the true population median WTP while the estimates from the more inclusive sample with 670 observations as an upper bound on the true median population WTP.

---

[14] We write $\hat{s} = -\hat{a}$ because we expect the propensity to purchase E85 to decline with respect to the price of E85 relative to E10 such that $\hat{a} < 0$. However, the scale parameter of a logistic distribution is positive, so we must have $s > 0$. To obtain a positive scale parameter $\hat{a}$, we add negative signs in the parametrization. Alternatively, we could have defined the ratio as the price of E10 divided by the price of E85 to find a positive value for $\hat{a}$.

[15] We also calculated the standard errors in the RP-data-only model by bootstrap, and the bootstrap standard errors are smaller than the standard errors from using the delta-method. However, we do not report bootstrap standard errors because their calculation is too computationally intensive in the other models.

We present estimates of the distribution parameters in Table 6. We express the price of E85 in E10-energy-equivalent terms, so a scale parameter of one would indicate that the median motorist values E85 and E10 the same in terms of cost per mile. A value smaller than one indicates that motorists discount E85 relative to E10 and a value greater than one indicates that motorists are willing to pay a premium for E85 relative to E10.

At the lower bound (sample with 479 observations), the median WTP for E85 is given by a price ratio between 0.65 and 0.55 at Retailer Y and a price ratio between 0.88 and 0.77 at Retailer Z, indicating that the median motorist discounts E85 relative to E10 at both retailers. The distribution of WTP is wide. In Model A, with a scale parameter of 0.654 and a shape parameter is 7.84, about 3.5 percent of motorists use E85 when the price ratio is one. Estimates of the median WTP are not statistically different across models with the 479-observation estimation sample. Adding probability weights to the likelihood function in Model C to correct for endogenous stratification has a small effect on the coefficient estimates compared to Model B. Adding the stated preference data does slightly improve precision for the median WTP coefficient estimates and significantly reduces the standard error for the shape coefficient estimates.

At the upper bound (sample with 670 observations), the median WTP for E85 is given by a price ratio between 0.81 and 0.71 at Retailer Y and at a price ratio between 1.48 and 1.30 at Retailer Z. Even at the upper bound, the median motorist at Retailer Y significantly discounts E85 below the cost-per-mile parity. In contrast, at Retailer Z the WTP of the median motorist is greater than one, indicating that if the (energy-adjusted) price of E85 equals the price of E10, the median flex motorist chooses E85.

Note that the median values for WTP we find are outside the estimation samples for the RP data model and at the edge of the estimation sample for the SP-off-RP models. We collected survey data to solve a problem encountered by Anderson (2012) who was only able to estimate a tail of the WTP distribution because the data only contained E85 prices higher than the E10 price when measured on the cost per mile basis. However, our results suggest that preferences are quite diffuse, and we should have generated even lower prices for E85 in our experiment to obtain better estimates of the lower tail of the distribution of WTP. The mean WTP we find is indeed lower than we anticipated when designing the survey. The scale and the shape parameters we report in Table 6 provide a good estimate of the upper tail of the distribution of WTP. However, we must be cautious when making inferences for values of the price ratio near or lower than the mean WTPs we report in Table 6.

Figure 4 shows the WTP distributions for Models A and B for the two retailers and the two estimation samples. We omit the estimated distribution of WTP from Model C from the figure in the

interest of clarity because it is similar to Model B. These cumulative distributions are analogous to demand curves for E85 where the price is the ratio of the energy-adjusted E85 and E10 prices and the quantity is the share of motorists who purchase E85.

For a given pair of retailer and estimation sample, the purchase probabilities for the two models are not statistically different from each other at a 95 percent confidence interval, in part because of the wide confidence intervals for Model A. Using the SP-off-RP approach provides a large gain in precision, especially when moving away from the observed average price ratio in the data of 1.09.

Table 6 and Figure 4 show that augmenting the RP data with the SP data reduces the size of the standard errors for the estimated parameters of the distribution of WTP for E85. The SP-off-RP approach uses data from an experiment where we presented a hypothetical price scenario to motorists. There are two issues we must consider: 1) whether there is hypothetical bias from the SP data, and 2) whether the hypothetical bias is a bigger issue than the problem of not observing revealed choices where the price of E85 is more favorable to its consumption. We do not believe that there is a significant hypothetical bias in our SP data. Conducting the experiment on-site minimizes biases that are typically associated with laboratory settings. Moreover, immediately before the survey began, motorists made an actual fuel purchase decision. The hypothetical fuel choices we offered were the same as the real fuel choices in every way except that their prices were changed, in a setting close to the actual purchase decision. For these reasons, we believe it is likely that the gain from improving the precision of our estimates by augmenting the RP data with SP data outweighs the hypothetical bias from the SP data in our study.

Our finding that motorists discount E85 relative to E10 on a cost per mile basis differs from previous studies that relied on stated preference data only (i.e., Jensen et. al. 2010; Petrolia et. al. 2010; Aguilar et. al. 2015). One possible explanation for the difference is that our data are not subject to the same hypothetical biases as the data collected in mail or online surveys. Another explanation is that our survey specifically targeted motorists with FFVs. Flex motorists may be more knowledgeable of E85 or have preferences different from the general population targeted in previous studies.

Summary and Conclusion

In this study, we conducted a survey of motorists at retail fuel stations offering E85 to estimate the distribution of preferences for E85 relative to E10 among flex motorists. Knowledge of US motorists' preferences for E85 is crucial for the evaluation of the RFS given that the implied mandated ethanol volumes now exceed the volumes that can easily be blended in regular gasoline. The EPA and other stakeholders expect that increased consumption of E85 will make compliance with the ethanol blending

mandates possible. The estimates of motorists' preferences for E85 relative to E10 in this study can be used to better measure the cost of compliance for increased volumes of biofuels in US motor fuels.

With the collaboration of two E85 retailers, we conducted an intercept survey at E85 stations to collect both revealed fuel preferences and stated fuel opinions from motorists with FFVs. We visited E85 stations in the urban areas of Colorado Springs, Des Moines, Little Rock, Tulsa, Los Angeles and Sacramento. We collected both RP and SP choice data giving us a greater range of relative fuel prices and hence more precise estimation of preferences. We specifically designed the experiment to induce switching and obtain a better mapping of motorists WTP for E85 relative to E10. We combined the RP and the SP data together to add variation in the fuel prices and to obtain more precise estimates of WTP for E85 at lower E85 prices.

We find stronger preferences for E85 in California than in the other states covered by our survey. When the nominal E85 price per gallon was about 80 percent of the nominal E10 price per gallon, less than half of flex motorists outside of California chose E85, whereas nearly 90 percent of flex motorists in California chose E85. Stronger self-selection of E85 motorists by California motorists because there are fewer E85 stations in California than in the other surveyed states partially explains the difference. After correcting for self-selection, our results indicate that motorists in the Los Angeles and Sacramento areas either have a genuine greater WTP for E85 as a substitute for E10 or that the California retailer's marketing techniques to promote biofuels to local flex motorists have been successful. One of the stations had only one E85 nozzle, and sometimes lines of FFVs would form to wait to refuel with E85 while E10 nozzles were available.

In the four states excluding California, we find a median WTP for the energy-adjusted E85 price between 55 and 76 percent of the price of E10, a significant discount. In California, we find a median WTP for the energy-adjusted E85 price between 77 and 131 percent of the price of E10. Estimates from the SP-off-RP models are similar to the estimates from the RP-only models, but the standard errors tend to be smaller because the SP data feature greater variation in 'observed' fuel prices.

In addition to the models we estimate where motorists respond the relative difference in fuel prices (the price ratio), we also estimate models described in a supplementary appendix online where motorists respond to the absolute difference in fuel prices (the E85 price premium). We find practically no difference in how well these two models fit the data. Thus, we cannot say whether motorists are responding to the E85 ratio or the E85 premium when they make their fuel choices.

We do find evidence for an unexplained difference in WTP between the motorists in California and those in other states. However, all else equal, the probability that a motorist chooses E85 is not

significantly different in Des Moines than it is in Colorado Springs, Little Rock, or Tulsa. This result holds even though general opinions about ethanol differ across states. Extrapolating to other regions of the United States, this result indicates that we may be able to apply estimation results from one state to project national demand, though we would need to make adjustments for California. Hopefully, external validity of our results will be confirmed in future research.

References

Adamowicz, W., J. Louviere, and M. Williams. 1994. Combining revealed and stated preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management,* 26(3), 271-292.

Aguilar, F., Z. Cai, P. Mohebalian, and W. Thompson. 2015. Exploring the drivers' side of the "blend wall". US consumer preferences for ethanol blend fuels. *Energy Economics,* 49, 217-226.

Alternative Fuels Data Center. 2017. Alternative Fuel and Advanced Vehicle Search. Available at: https://www.afdc.energy.gov/vehicles/search/.

Anderson, S. T. and J. M. Sallee. 2011. Using loopholes to reveal the marginal cost of regulation: The case of fuel-economy standards. *The American Economic Review*, 101(4), 1375-1409.

Anderson, S. 2012. The demand for ethanol as a gasoline substitute. *Journal of Environmental Economics and Management,* 63(2), 151-168.

Ben-Akiva, M. and T. Morikawa. 1990. Estimation of switching models from revealed preferences and stated intentions. *Transportation Research Part A: General,* 24(6), 485-495.

Bracmort, K. 2018. The Renewable Fuel Standard (RFS): An Overview. CRS Report No. R43325. Washington, DC: Congressional Research Service. Available at https://fas.org/sgp/crs/misc/R43325.pdf.

Clinton, B., C. Johnson, K. Moriarty, E. Newes, and L. Vimmerstedt. 2017. Preliminary assessment of spatial competition in the market for E85: Presentation supplement. Report no. NREL/TP-6A20-66943. National Renewable Energy Laboratory (NREL). Available at http://www.nrel.gov/docs/fy17osti/66943.pdf.

Corts, K. 2010. Building out alternative fuel retail infrastructure: Government fleet spillovers in E85. *Journal of Environmental Economics and Management,* 59(3), 219-234.

Domencich, T. and D. McFadden. 1975. Urban Travel Demand - A Behavioral Analysis. North-Holland Publishing Company. Available at https://eml.berkeley.edu/~mcfadden/travel.html.

Hensher, D. and M. Bradley. 1993. Using stated response choice data to enrich revealed preference discrete choice models. *Marketing Letters,* 4(2), 139-151.

Hensher, D., J. Louviere, and J. Swait. 1999. Combining sources of preference data. *Journal of Econometrics,* 89(1), 197-221.

Houde, J. 2012. Spatial differentiation and vertical mergers in retail markets for gasoline. *The American Economic Review,* 102(5), 2147-2182.

Imbens, G.W., 1992. An efficient method of moments estimator for discrete choice models with choice-based sampling. *Econometrica*, 60(5): 1187-1214.

Imbens, G.W. and T. Lancaster. 1996. Efficient estimation and stratified sampling. *Journal of Econometrics,* 74(2), 289-318.

Jensen, K., C. Clark, B. English, R. Menard, D. Skahan, and A. Marra. 2010. Willingness to pay for E85 from corn, switchgrass, and wood residues. *Energy Economics,* 32(6), 1253-1262.

King, G. and L. Zeng., 2001. Logistic regression in rare events data. *Political Analysis*, 9(2), 137-163.

Knittel, C.R., B,S, Meiselman, and J.H. Stock, 2017. The pass-through of RIN prices to wholesale and retail fuels under the Renewable Fuel Standard. *Journal of the Association of Environmental and Resource Economists*, 4(4): 1081-1119.

Manski, C. and S. Lerman. 1977. The estimation of choice probabilities from choice based samples. *Econometrica*, 45(8): 1977-1988.

Manski, C.F. and D. McFadden. 1981. Alternative estimators and sample designs for discrete choice analysis. In C. F. Manski, & D. McFadden (Eds.), *Structural Analysis of Discrete Data with Econometric Applications* Cambridge: The MIT Press, pp. 2-50.

Petrolia, D., S. Bhattacharjee, D. Hudson, and C. Herndon. 2010. Do Americans want ethanol? A comparative contingent-valuation study of willingness to pay for E-10 and E-85. *Energy Economics,* 32(1), 121-128.

Pouliot, S. and B. A. Babcock. 2014. The demand for E85: Geographical location and retail capacity constraints. *Energy Economics*, 45, 134-143.

Pouliot, S., and B. A. Babcock. 2016. Compliance path and impact of ethanol mandates on retail fuel market in the short run. *American Journal of Agricultural Economics,* 98(3), 744-764.

Rubin, J. 1996. A model of intertemporal emission trading, banking and borrowing. *Journal of Environmental Economics and Management,* 31(3), 269-286.

Rubin, J. and P. Leiby. 2000. An analysis of alternative fuel credit provisions of US automotive fuel economy standards. *Energy Policy*, 28(9), 589-601.

Salvo, A. and C. Huse. 2013. Build it, but will they come? Evidence from consumer choice between gasoline and sugarcane ethanol. *Journal of Environmental Economics and Management,* 66(2), 251-179.

Train, K. and W. Wilson. 2008. Estimation on stated-preference experiments constructed from revealed-preference choices. *Transportation Research Part B: Methodological,* 42(3), 191-203.

Train, K. and W. Wilson. 2009. Monte Carlo analysis of SP-off-RP data. *Journal of Choice Modelling,* 2(1), 101-117.

U.S. Energy Information Administration. 2017. Annual Energy Outlook 2017. Available at https://www.eia.gov/outlooks/archive/aeo17/.

U.S. Energy Information Administration. 2018a "Short-Term Energy Outlook: Table 4a. U.S. Petroleum and Other Liquids Supply, Consumption and Inventories."
Available at https://www.eia.gov/outlooks/steo/data/browser/#?v=9.

U.S. Energy Information Administration. 2018b Alternative Fuel Vehicle Data: Supplier data. Available at https://www.eia.gov/renewable/afv/supply.php.

Wolff, H. 2014. Value of time: Speeding behavior and gasoline prices. *Journal of Environmental Economics and Management*, 67(1), 71-88.

Tables

Table 1. Summary of fuel prices, fuel choices and questions for sample selection

Location	# obs.	Avg. E85 price	Avg. E10 price	Avg. ratio	Min. ratio	Max. ratio	% E85	% of E85 who did not drive	% of E85 with dist. = 0
Col. Springs 1	11	2.57	2.00	1.28	1.28	1.28	9.1	0.0	0.0
Col. Springs 2	33	2.57	2.02	1.27	1.25	1.28	30.3	30.0	90.0
Col. Springs 3	54	2.57	2.06	1.25	1.25	1.25	13.0	0.0	42.9
Total	**98**	**2.57**	**2.04**	**1.26**	**1.25**	**1.28**	**18.4**	**16.7**	**66.7**
Des Moines 1	117	2.77	2.72	1.02	0.94	1.18	46.2	3.7	51.9
Des Moines 2	50	2.95	2.64	1.12	0.96	1.19	28.0	14.3	64.3
Des Moines 3	27	2.97	2.81	1.06	0.96	1.12	51.9	0.0	14.3
Des Moines 4	114	2.94	2.69	1.10	0.96	1.19	40.4	10.9	52.2
Total	**308**	**2.88**	**2.70**	**1.07**	**0.94**	**1.19**	**41.6**	**7.0**	**49.2**
Little Rock 1	26	2.36	2.18	1.08	1.06	1.11	34.6	11.1	66.7
Little Rock 2	23	2.35	2.13	1.10	1.10	1.10	34.9	0.0	50.0
Little Rock 3	60	2.35	2.18	1.08	1.08	1.08	31.7	10.5	52.6
Total	**109**	**2.35**	**2.17**	**1.12**	**1.06**	**1.11**	**33.0**	**8.3**	**55.6**
Tulsa 1	58	2.31	2.09	1.10	1.10	1.11	41.4	8.3	50.0
Tulsa 2	12	2.31	2.10	1.10	1.10	1.10	66.7	12.5	50.0
Tulsa 3	65	2.31	2.04	1.13	1.13	1.13	18.5	8.3	58.3
Total	**135**	**2.07**	**2.07**	**1.12**	**1.10**	**1.13**	**32.1**	**9.1**	**52.3**
Retailer Y total	**650**	**2.63**	**2.38**	**1.10**	**0.94**	**1.28**	**34.8**	**8.4**	**52.2**
Los Angeles 1	85	3.35	3.20	1.05	1.04	1.06	95.3	3.7	39.5
Los Angeles 2	52	3.38	3.10	1.09	1.08	1.10	84.6	2.3	50.0
Total	**137**	**3.36**	**3.16**	**1.06**	**1.04**	**1.10**	**90.5**	**3.2**	**43.2**
Sacramento 1	43	3.29	3.23	1.02	1.01	1.02	81.4	2.9	54.3
Sacramento 2	51	3.19	2.92	1.09	1.08	1.09	88.2	11.1	64.4
Total	**94**	**3.24**	**3.06**	**1.06**	**1.01**	**1.09**	**85.1**	**7.5**	**60.0**
Retailer Z total	**231**	**3.31**	**3.12**	**1.06**	**1.01**	**1.10**	**88.7**	**4.9**	**49.8**
Sample total	**881**	**2.81**	**2.58**	**1.09**	**0.94**	**1.28**	**48.9**	**6.7**	**51.0**

Note: All prices are in dollars per gallon. Prices for E85 are in energy-adjusted terms relative to E10. The E85 ratio is the E85 price divided by the E10 price. The column "% E85" is the share of motorists who refueled with E85. The column "% of E85 who did not drive" is the share of E85 motorists who responded that they did not drive to the fuel station because it offered E85. "% of E85 with dist. = 0" is the share of E85 motorists who either did not drive to the fuel station because it offered E85 or indicated that they did not drive out of their way at all to access the fuel station.

Table 2: Responses to questions to flex motorists who refueled with E10

	Yes	No / Don't know	Total
Is your vehicle an FFV?	385 (86%)	65 (14%)	450
Have you ever fueled with E85?	145 (38%)	240 (62%)	385
Did you know this station sells E85?	280 (73%)	105 (27%)	385

Note: The complete intercept survey that includes these questions is available in a supplementary appendix online.

Table 3: Responses to fuel opinion questions by region and fuel choice

Which fuel...	State	E10 motorists (%)				E85 motorists (%)			
		E85	E10	No diff	Don't know	E85	E10	No diff	Don't know
is better for the environment?	CO	64	5	12	19	67	6	11	17
	IA	77	7	11	6	84	1	11	5
	AR	52	18	14	16	67	6	19	8
	OK	42	19	25	14	43	9	30	18
	CA	58	15	8	19	83	0	7	10
	Total	**62**	**11**	**14**	**12**	**77**	**2**	**12**	**10**
is better for your engine?	CO	24	48	15	14	39	33	6	22
	IA	25	42	18	14	41	31	17	10
	AR	18	52	14	16	44	28	17	11
	OK	13	69	9	9	32	34	16	18
	CA	8	42	15	35	60	10	14	15
	Total	**20**	**50**	**15**	**15**	**50**	**21**	**15**	**14**
is better for the economy?	CO	44	30	11	15	44	33	17	6
	IA	71	12	10	8	88	4	4	5
	AR	33	45	4	18	56	19	17	8
	OK	25	44	15	15	55	20	16	9
	CA	35	31	8	27	73	12	8	7
	Total	**48**	**28**	**10**	**13**	**73**	**12**	**9**	**7**
is better for national security?	CO	34	31	24	11	28	11	28	33
	IA	51	13	11	25	70	7	5	17
	AR	32	29	12	27	44	28	11	17
	OK	21	29	30	21	30	23	20	27
	CA	23	27	8	42	44	9	9	38
	Total	**37**	**23**	**17**	**23**	**50**	**11**	**10**	**29**
yields more miles per gallon?	CO	16	60	5	19	17	61	6	17
	IA	8	76	4	12	20	66	5	10
	AR	12	66	7	15	28	56	8	8
	OK	10	67	3	20	23	52	5	20
	CA	8	50	0	42	24	46	15	15
	Total	**10**	**68**	**4**	**17**	**23**	**54**	**10**	**13**

Note: A supplementary appendix online shows the complete questionnaire and specific framing of the questions.

Table 4. Mean of marginal effects for models with 479 observations

	A) RP data		B) SP-off-RP		C) Endo strat	
	Marg. Eff.	Std. Err.	Marg. Eff.	Std. Err.	Marg. Eff.	Std. Err.
Log (ratio)	-0.342**	0.170	-0.249**	0.057	-0.242**	0.041
Government vehicle	0.111**	0.058	0.084	0.072	0.086	0.060
Company vehicle	-0.061	0.039	-0.029	0.028	-0.028	0.025
Other owner	0.038	0.033	0.014	0.037	-0.015	0.037
Truck	-0.009	0.028	-0.022	0.020	-0.018	0.017
SUV	-0.007	0.028	-0.013	0.021	-0.020	0.018
Van	0.019	0.029	-0.004	0.031	-0.009	0.026
FFV badge	-0.030	0.020	0.004	0.021	0.013	0.017
Female	0.009	0.022	-0.044	0.023	-0.036*	0.019
Log age	-0.034	0.030	-0.041	0.026	-0.009	0.022
Log miles per year	-0.019	0.017	-0.042**	0.014	-0.031**	0.012
Eth. better for env.	-0.017	0.040	-0.017	0.034	-0.017	0.028
Gas better for env.	-0.082	0.059	-0.093*	0.050	-0.094**	0.047
No diff. for env.	-0.098*	0.058	-0.095**	0.044	-0.087**	0.038
Eth. better for eng.	0.080**	0.038	0.045	0.034	0.052	0.032
Gas better for eng.	0.049	0.039	0.023	0.031	0.037	0.029
No diff. for eng.	-0.016	0.057	0.000	0.034	0.029	0.030
Eth. better for econ.	0.084	0.053	0.103**	0.049	0.080**	0.038
Gas better for econ.	0.025	0.056	0.067	0.050	0.039	0.038
No diff. for econ.	0.132**	0.057	0.130**	0.057	0.096**	0.044
Eth. better for nat. sec.	0.014	0.026	0.014	0.024	0.004	0.020
Gas better for nat. sec.	0.040	0.031	0.030	0.028	0.025	0.024
No diff. for nat. sec.	-0.053	0.042	0.002	0.029	-0.001	0.024
Eth. yields better mpg	0.009	0.039	-0.025	0.040	-0.019	0.036
Gas yields better mpg	0.036	0.033	0.015	0.028	0.018	0.025
No diff. for mpg	0.059	0.052	0.079**	0.035	0.069**	0.031
Arkansas	0.025	0.034	0.008	0.028	0.008	0.024
Colorado	0.055	0.048	0.019	0.027	0.026	0.023
Oklahoma	0.062	0.035	0.027	0.025	0.023	0.022
California	0.134**	0.029	0.097**	0.029	0.086**	0.025

* Significant at 90 percent; ** Significant at 95 percent. Sample size is 479 observations using the stricter sample-selection rule. For categorical variables, the reference categories are personal vehicle for ownership (Gov., Company and Other owner), car for vehicle type (Truck, SUV and Van), no FFV badge for the presence of a badge on the vehicle (Badge), "Don't know" for all opinion questions and the State of Iowa for the States (Arkansas, Colorado, Oklahoma and California).

Table 5. Mean of marginal effects for models with 670 observations

	A) RP data		B) SP-off-RP		C) Endo strat	
	Marg. Eff.	Std. Err.	Marg. Eff.	Std. Err.	Marg. Eff.	Std. Err.
Log (ratio)	-0.681**	0.212	-0.616**	0.056	-0.546**	0.061
Government vehicle	0.387**	0.078	0.470**	0.107	0.441**	0.097
Company vehicle	-0.033	0.050	-0.018	0.048	-0.007	0.045
Other owner	-0.022	0.067	-0.064	0.071	-0.045	0.066
Truck	0.040	0.038	0.018	0.034	0.014	0.032
SUV	-0.052	0.040	-0.049	0.034	-0.045	0.032
Van	-0.036	0.048	-0.050	0.048	-0.058	0.046
FFV badge	0.034	0.031	0.068**	0.029	0.070**	0.027
Female	0.039	0.034	0.010	0.033	0.005	0.031
Log age	0.049	0.043	-0.003	0.042	0.007	0.039
Log miles per year	-0.025	0.023	-0.015	0.020	-0.018	0.019
Eth. better for env.	-0.023	0.053	-0.018	0.045	0.002	0.044
Gas better for env.	-0.141	0.085	-0.154*	0.080	-0.114	0.073
No diff. for env.	-0.038	0.064	-0.040	0.057	-0.024	0.055
Eth. better for eng.	0.129**	0.046	0.106**	0.043	0.094**	0.041
Gas better for eng.	-0.045	0.048	-0.049	0.043	-0.052	0.041
No diff. for eng.	0.038	0.052	0.032	0.045	0.023	0.043
Eth. better for econ.	0.162**	0.055	0.147**	0.050	0.145**	0.048
Gas better for econ.	0.022	0.065	0.040	0.057	0.043	0.054
No diff. for econ.	0.145	0.070	0.079	0.061	0.080	0.059
Eth. better for nat. sec.	0.035	0.037	0.039	0.034	0.042	0.033
Gas better for nat. sec.	0.046	0.051	0.027	0.048	0.025	0.045
No diff. for nat. sec.	-0.028	0.053	0.022	0.042	0.041	0.040
Eth. yields better mpg	0.037	0.055	0.013	0.054	0.020	0.051
Gas yields better mpg	0.095**	0.044	0.067	0.043	0.080*	0.041
No diff. for mpg	0.157**	0.069	0.122**	0.053	0.149**	0.051
Arkansas	0.010	0.046	-0.023	0.048	-0.015	0.045
Colorado	0.026	0.065	-0.011	0.049	0.012	0.044
Oklahoma	0.091**	0.046	0.043	0.041	0.064*	0.038
California	0.414**	0.034	0.321**	0.030	0.329**	0.028

* Significant at 90 percent; ** Significant at 95 percent. Sample size is 670 observations using the looser sample-selection rule. For categorical variables, the reference categories are personal vehicle for ownership (Gov., Company and Other owner), car for vehicle type (Truck, SUV and Van), no FFV badge for the presence of a badge on the vehicle (Badge), "Don't know" for all opinion questions and the State of Iowa for the States (Arkansas, Colorado, Oklahoma and California).

**Table 6.** Estimated WTP distribution parameters

1) Lower Bound – Sample with 479 observations

	A) RP data		B) SP-off-RP		C) Endo strat	
	Estimates	Std. Err.	Estimates	Std. Err.	Estimates	Std. Err.
Scale at Retailer Y	0.654**	0.217	0.554**	0.167	0.587**	0.146
Scale at Retailer Z	0.875**	0.243	0.769**	0.264	0.775**	0.230
Shape	7.844**	3.831	5.504**	1.055	6.649**	1.053

2) Upper Bound – Sample with 670 observations

	A) RP data		B) SP-off-RP		C) Endo strat	
	Estimates	Std. Err.	Estimates	Std. Err.	Estimates	Std. Err.
Scale at Retailer Y	0.807**	0.152	0.756**	0.187	0.707**	0.188
Scale at Retailer Z	1.482**	0.351	1.311**	0.288	1.300**	0.317
Shape	4.990**	1.588	4.487**	0.386	4.249**	0.429

* Significant at 90 percent; ** Significant at 95 percent. The model samples are: A) MLE with the RP data only; B) SP-off-RP method for data augmentation; C) SP-off-RP method with weights to account for endogenous stratification. The scale coefficients of the log-logistic distribution are the median of the distribution function.
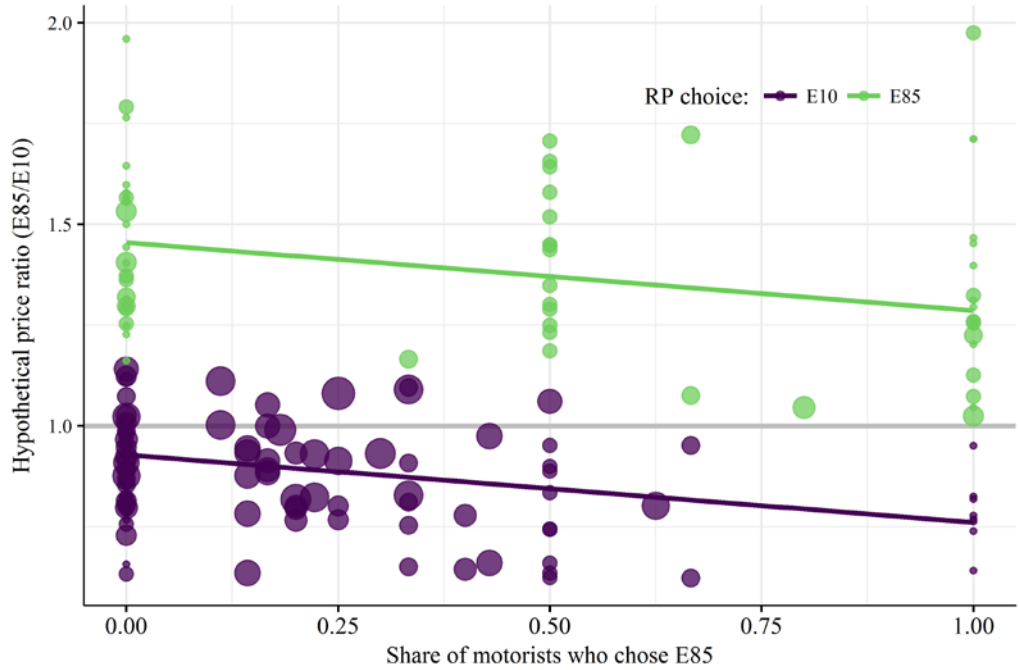
Figures
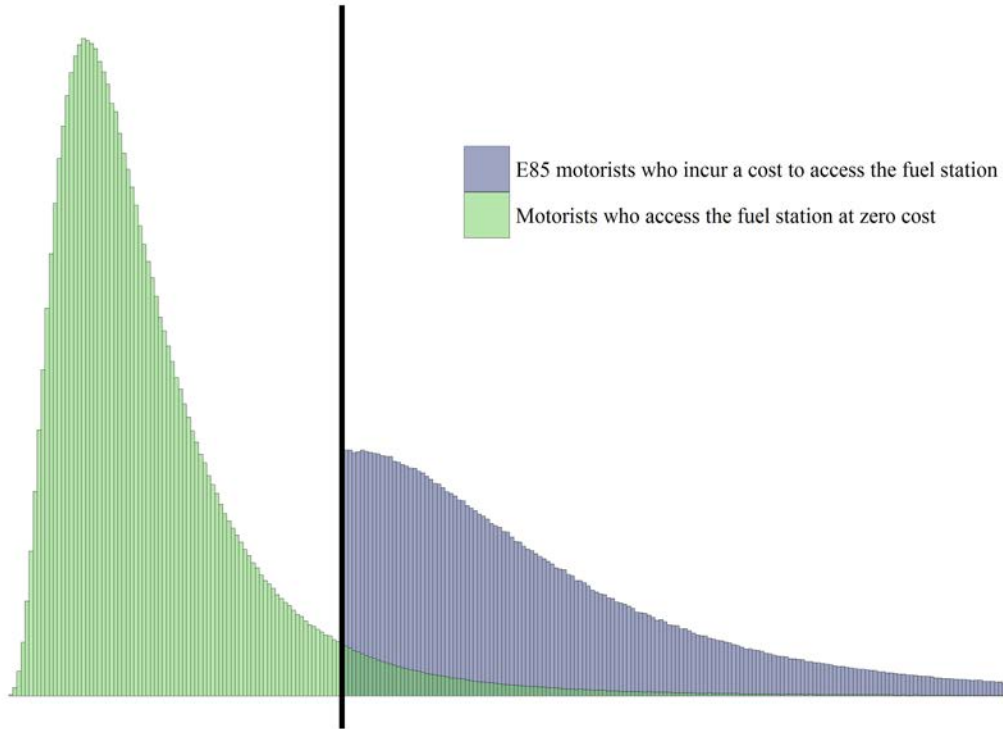
Figure 1. Self-selection at a station that offers E85

Note: The vertical black line is the station's posted net E85 price ratio and cutoff point where motorists on the left purchase E10 and motorists to the right purchase E85. The areas for the two categories of motorists represents the number of motorists in each category. The area in green is for motorists who can access the fuel station at zero costs, for which it is possible to sample from the entire population at a fuel station. The area in purple is for motorists who patronize the fuel station specifically because it offers E85, and E85 is the only choice observable at the fuel station. These motorists must incur a cost to access the fuel station.
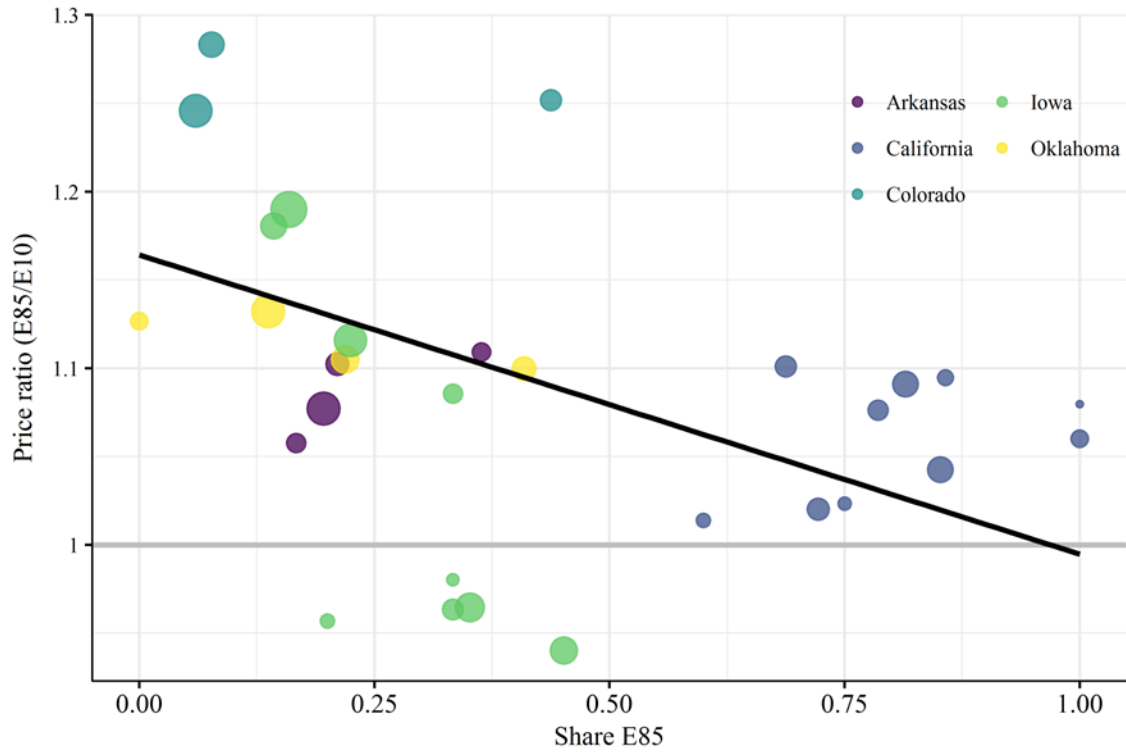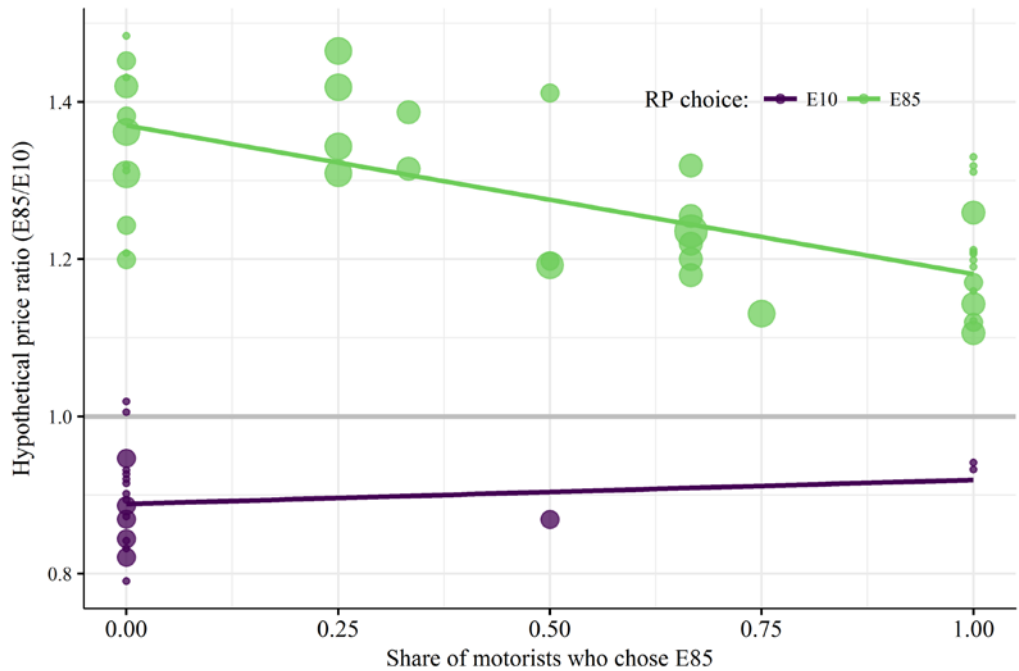
Figure 2. Price ratio and share of motorists who chose E85 in the sample with 670 observations

Note: E85 prices are in E10-energy-equivalent terms. The sizes of the dots represent the relative number of observations at each pair of price ratio and share of motorists. The line is a weighted linear regression.

a) Hypothetical E85 price ratio and share of motorists who chose E85 at Retailer Y in the sample with 670 observations

b) Hypothetical E85 price ratio and share of motorists who chose E85 at Retailer Z in the sample with 670 observations

Figure 3. Hypothetical price ratio and share of motorists who chose E85 in the sample with 670 observations

Note: E85 prices are in E10-energy-equivalent terms. The sizes of the dots represent the relative number of observations at each pair of price ratio and share of motorists. The lines are weighted linear regressions.
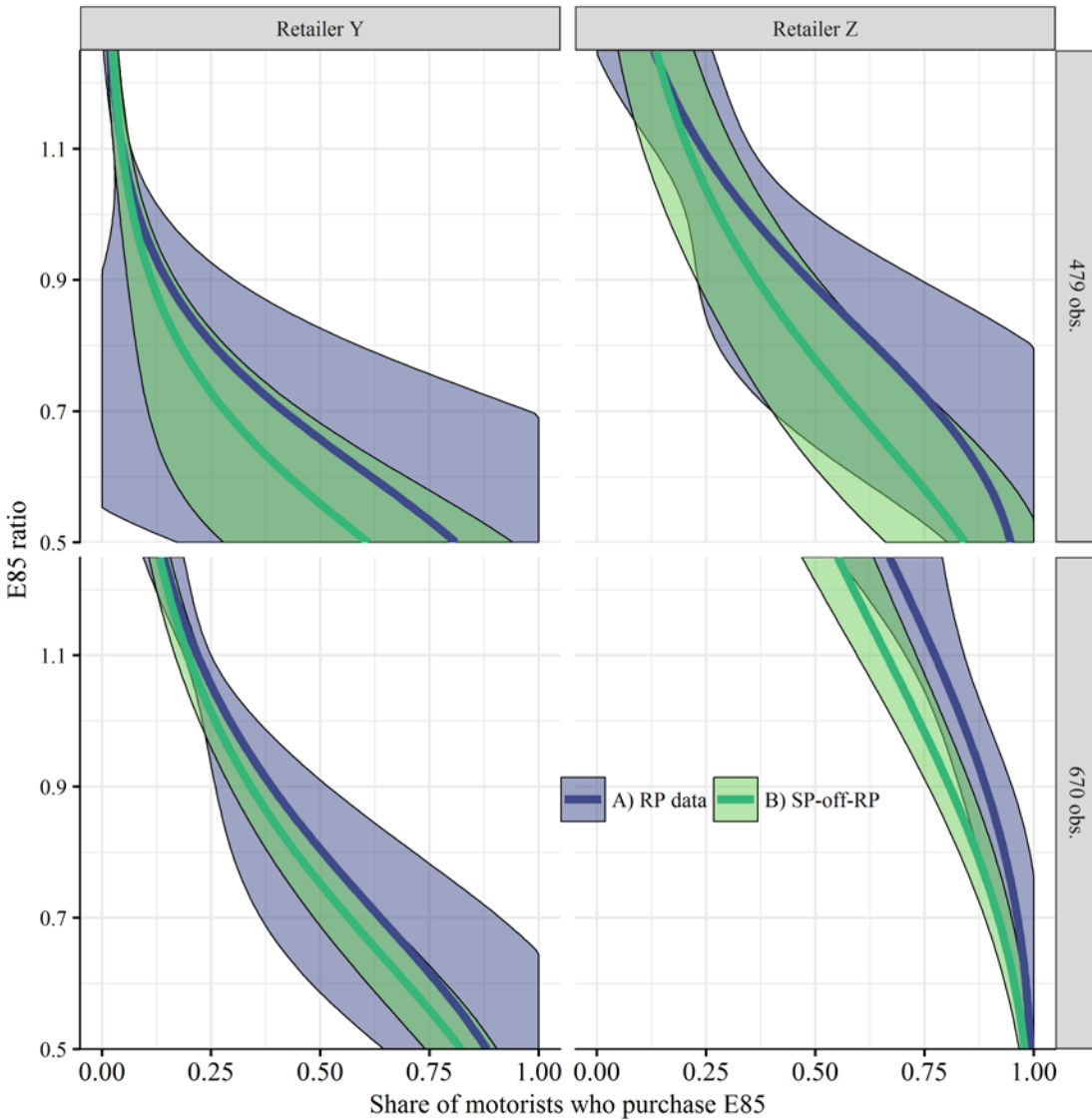
Figure 4. Estimated probability of E85 purchase

Note: The price ratio is the price of E85 divided by the price of E10 with E85 prices measured in E10-energy-equivalent terms, so an E85 ratio of one equates cost per mile of driving. The shaded areas are 95-percent confidence intervals.