Econometric Analysis of Motorists' Preference for Ethanol in Motor Fuel

Kenneth Liao

PhD candidate

Department of Economics

Iowa State University

kliao@iastate.edu

Sebastien Pouliot

Associate professor

Department of Economics

Iowa State University

pouliot@iastate.edu

Selected Paper prepared for presentation at the 2015 Agricultural & Applied Economics Association and Western Agricultural Economics Association Annual Meeting, San Francisco, CA, July 26-28

Introduction

The second installment of the Renewable Fuels Standard (RFS2) requires minimum blending of ethanol and other biofuels into the motor fuel consumed in the United States. The vast majority of gasoline consumed in the United States contains no more than 10 percent ethanol. This gasoline-ethanol blend is conventionally known as E10. The maximum quantity of ethanol that can be blended into the total motor fuel pool through E10 is commonly referred to as the E10 blend wall. The quantity of ethanol mandated by the RFS2 is now reaching the point where it is set to surpass the E10 blend wall.

One solution to the blend wall is the consumption of gasoline blends that contain more than 10 percent ethanol such as E85, which contains no more than 83 and no less than 51 percent ethanol. On average, a gallon of E85 contains about 74 percent ethanol so each gallon of E85 consumed as a substitute for E10 increases aggregate ethanol consumption by about 0.64 gallons (EIA 2015). As such, ethanol consumption can exceed the blend wall if some motorists fuel with E85 instead of E10. However, E85 consumption in the United States has historically been limited, and it is not at the level needed to meet the expanded ethanol mandates.

The RFS2 provides for minimum quantities of biofuels and a credit system called Renewable Identification Number (RIN) that creates a tax on gasoline and a subsidy to biofuels. The size of the tax and the size of the subsidy endogenously adjust to cost of production, the strength of the demand for biofuels, and mandated volumes. Therefore analysis of the policy requires a description of the demand curve for biofuels and in this particular case the demand for ethanol beyond the E10 blend wall. This paper empirically estimates the relative preferences of motorists for E10 and E85, and provides insight to the question, "Who will consume the additional ethanol, and at what price?". We use market data to estimate willingness to pay for E85 relative to E10. Estimates of motorists' willingness to pay to use E85 instead of E10 can be used to understand the feasibility of expanding the consumption of ethanol (e.g., Babcock & Pouliot (2013a)). Our study allows prediction of the share of flex motorists who choose E85 instead of E10 given fuel prices, and our results can be used to evaluate the welfare impacts of the biofuels mandate (e.g., Anderson (2012), Babcock & Pouliot (2013b)).

The demand for E85 is limited to motorists who drive flexible-fuel vehicles (FFVs). Unlike drivers of conventional vehicles, motorists with FFVs, which we refer to as 'flex motorists', can choose to fill with any gasoline blend that contains between 0 and 85 percent ethanol. E85 yields lower fuel economy than E10 because ethanol has lower energy content per volume than gasoline. Even so, some motorists may choose E85 when its price (in energy-equivalent terms) is at a premium relative to E10. On the other hand, some motorists may choose not to fuel with E85 even when its energy-adjusted price is at a discount relative to E10.

Efforts have been made to understand the demand for E85 in the United States, but they have been somewhat limited due to the lack of available E85 data. Anderson (2012) estimates willingness to pay for E85 using E85 sales volume and price data from retail stations in Minnesota between 1997 and 2006. During that time period, the energy-adjusted price of E85 was almost always greater than the price of E10. As a result, Anderson (2012) is unable to recover the full distribution of consumer willingness to pay for E85 and instead estimates the upper tail of a distribution where the energy-adjusted price of E85 is higher than the price of E10, and only flex motorists with high willingness to pay for E85 use it.

Other recent studies that use Minnesota data to study E85 demand include Corts (2010) and Liu and Greene (2013). Corts (2010) recognizes that most of the early data represent E85 use by government fleet vehicles. The study attempts to test that government fleet FFV mandates encourage retail fuel stations to invest in E85 fueling infrastructure and that increased availability of E85 increases consumer demand for FFVs. Corts (2010) shows that government fleet adoption of FFVs led to an increase in the number of retail E85 stations, but concedes that he is unable to test the second hypothesis due to the limitations of the data. Specifically, Corts (2010) notes that most FFVs in his dataset were purchased prior to the widespread availability of E85, and many of the owners of these vehicles may not even know of the vehicles' capabilities. Corts (2010) concludes that data from more recent years is required to estimate a credible model of retail E85 and FFV demand.

Liu and Greene (2013) estimate E85 demand using more recent data which allows a better estimate of non-fleet demand than previous studies. The dependent variable is the share of energy services provided to flex motorists in Minnesota that is attributable to ethanol, and Liu and Greene (2013) find a high price elasticity of demand for E85.

This study expands on the work of Anderson (2012). In recent years, the number of fuel stations that offer E85 has increased, E85 prices have fallen relative to E10 prices so that E85 is sometimes offered at a discount, and the majority of E85 sales are now to private motorists rather than government fleet vehicles. Recent data on E85 sales and prices covering a wider range of E85-E10 price differences allow us to trace more completely and more precisely the distribution of willingness to pay for E85 as a substitute for E10 among motorists with FFVs. Our complete dataset consists of over 21,000 monthly observations of E85 sales volumes and volume-weighted prices from over 400 retail fuel stations in Minnesota from between October 1997 and August 2014.

We find that the average flex motorist discounts E85 by about \$0.57 per gallon when measured in E10 energy-equivalent dollars. The mean E10 price for our data is about \$3.33, so the average flex motorist requires the E85 price to be about 17 percent lower than the energy-equivalent price to choose E85. If E85 were priced at parity with E10 in energy-equivalent dollars, we find that about 11 percent of flex motorists choose E85.

The next section contains background information on the E85 industry. We describe the theoretical model based on Anderson (2012) in section III. In section IV, we explain the empirical model. Section V describes the data. We present our estimation results in section VI. And lastly, section VII concludes.

E85 in Minnesota

To estimate relative preferences for E10 and E85, we consider the technical capabilities of FFVs and the markets in which the retail E85 stations operate.

FFVs and E85 Stations

Most traditional automobiles cannot accommodate gasoline blends with higher than 10 or 15 percent ethanol by volume. FFVs can operate using a range of gasoline blends including E10, E85, and any combination of the two. Because ethanol contains about 2/3 of the energy of gasoline, a flex-fuel vehicle running on E85 gets between 75 and 80 percent as many miles per gallon compared to E10, depending on the specific vehicle and the exact concentration of ethanol in the E85 fuel blend. To facilitate the comparison of E85 and E10, we convert E85 prices and volumes into their E10 energy equivalent.

Most FFVs are simply alternate versions of conventional vehicle models. Automobile manufacturers have incentives from the Corporate Average Fuel Economy (CAFE) standards to produce FFVs. Up to a certain annual limit, FFVs are treated as though they are operated partially on E85, but the fuel economy of FFVs is calculated as the total miles the vehicle can travel per gallon of gasoline input (the ethanol fuel input is excluded in the fuel economy calculation). For motorists, the operation of an FFV is identical to a conventional vehicle. Automobile manufacturers note that there is essentially no difference in performance for an FFV using E85 compared to using E10 other than the difference in fuel economy. In many cases, consumers are not able to acquire a certain vehicle make and model in anything but the FFV version or are unaware that they have purchased an FFV. Thus whether a motorist owns an FFV may be independent of ethanol preference and price (Corts 2010).

Most retail fuel stations do not supply E85 because it requires a dedicated underground storage tank and the pumps that dispense E85 require modifications to withstand the greater corrosive properties of ethanol. The cost to install new fueling infrastructure can be significant for retailers, who are understandably hesitant to make such an investment without knowing what E85 demand will be in the future.

E85 Data

There is no comprehensive source of data on national E85 sales or prices in the United States. Some recent studies have conducted surveys to obtain stated-preference data which are then used to estimate motorists' willingness to pay to use E85. See for example Jensen et al. (2010), Petrolia et al. (2010), and Aguilar et al. (2015). The best available revealed-preference data on E85 sales come from the Minnesota Department of Commerce (MN DoC) survey.[1] The state of Minnesota has been promoting ethanol production and use with supply-side incentives since the 1980s. As a result, E85's market share in Minnesota is relatively high compared to other states, fuel stations offering E85 are relatively abundant compared to other states, and a majority of sales are to private (non-fleet) motorists. Minnesota was the first state to require that nearly all gasoline blends contain at least 10 percent ethanol and has continued to provide incentives to ethanol producers, blenders, and retailers. Minnesota supplies retail fuel stations with government loans to pay for E85 infrastructure costs. Retailers can have these loans partially or completely forgiven by reporting E85 sales volumes and revenues in a monthly survey conducted by the Minnesota Department of Commerce. This survey is the primary source of the data we use in our estimation.

The available Minnesota data start in 1997, when only a handful of E85 stations were operating in the state, and E85 consumers were almost exclusively government fleet vehicles required by law to use E85 whenever possible. Figure 1 shows that from 1997 to 2004, the average monthly E85 sales volumes from E85 stations in Minnesota increased steadily from about 200 gallons to about 2,500 gallons. In 2005 and 2006 there was a large increase, and by 2006, the average monthly E85 sales volume had grown to about 7,000 gallons. Figure 1 also shows a seasonality effect in the E85 sales volumes; Minnesotans drive more in the summer, and so that is when more fuel is sold.

Along with the average monthly E85 sales volume per station, Figure 1 shows the total monthly consumption of E85 in Minnesota. Even though not all E85 stations report to MN DoC each month, MN

[1] The Iowa Department of Revenue also collects data on E85 consumption but the data are not as detailed as the survey data from MN DoC.

DoC keeps track of the total number of operating E85 stations, and the total monthly E85 consumption in Minnesota is calculated as the average E85 sales volume among reporting E85 stations multiplied by the total number of E85 stations operating in Minnesota that month.

Figure 1 shows that the monthly quantity of E85 sold in Minnesota grew steadily from fewer than 2,000 gallons in 1997 to about 250,000 gallons by 2004. Monthly E85 sales increased to over 800,000 gallons in the summer of 2005, and again to over 1,600,000 gallons in the summer of 2006. There is a noticeable decrease in total sales in 2007 likely due to the recession and a high ethanol price. Figure 1 also shows that total E85 sales in Minnesota fell significantly in 2012. That year, the United States experienced a drought that significantly reduced corn yields, causing corn prices to rise and making the ethanol that enters into E85 more expensive.

Figure 2 shows that from 1997 to 2004, the number of fuel stations that offered E85 in Minnesota grew steadily from fewer than 10 to about 100, and, like E85 sales volumes, in 2005 and 2006 the number of E85 stations increased significantly, so that there were about 300 E85 stations in Minnesota by the end of 2006. Figure 2 also shows that the growth in the number of retail E85 stations plateaued at around 350 in 2009, and there was a small drop in the number of E85 stations in Minnesota at the beginning of 2014. Because both average E85 sales per station and the number of E85 stations in Minnesota increase sharply between 2004 and 2006, the increase in total statewide E85 consumption in those years is even more prominent

Theoretical Model

We derive the demand for E85 based on a choice model described in Anderson (2012). The model is especially useful to formalize the connection between flex motorists' fuel preferences and aggregate market demand for E85. We provide below a brief description of the model and refer the interested reader to Anderson (2012) for additional details.

Motorist Behavior

Each motorist who owns an FFV maximizes the quasi-linear utility function

$$U = v\left(\left(q_e + q_g\right)m\right) + \theta_e q_e + \theta_g q_g + z, \tag{1}$$

where q_e is the quantity of E85 in gallons, q_g is the quantity of E10 in gallons, m is the fuel economy of the vehicle in miles per gallon, and z is a numeraire that captures the consumption of all other goods measured in dollars. The first term of the utility function represents the utility gained from driving M miles where $M \equiv \left(q_e + q_g\right)m$, and $v(M)$ is increasing and concave in miles driven. The quantity of E85 is expressed in E10 energy-equivalent gallons. As such, ethanol and gasoline are perfect substitutes in producing miles. The parameters θ_e and θ_g measure the utility from consuming one gallon of E85 or E10 respectively for attributes of the fuel other than its main function to provide vehicle miles.

The parameters θ_e and θ_g are motorist-specific and allow fuel choice to affect utility in a way that is unrelated to the cost per mile driven. This means that the motorists will not always choose the fuel with the lowest energy-adjusted price. Motorists receive some direct utility benefit or incur some direct utility cost from fuel consumption unrelated to the fuel's primary use providing energy for the vehicle. For example, some motorists may be willing to pay more for E85 because they value the environmental benefits of using renewable fuels while other motorists may be willing to pay more for E10 to avoid more frequent refueling.

Each motorist faces the budget constraint

$$p_e q_e + p_g q_g + z \leq y,$$

where p_g is the price of E10, p_e is the price of E85 (converted to E10 energy-equivalent dollars), y is the motorist's income, and the price of the composite good z is normalized to 1. By Walras' Law, the budget constraint holds with equality implying

$$z = y - p_e q_e - p_g q_g.$$

Substituting the value of z into equation (1), the unconstrained utility maximization problem for flex motorists is

$$\max_{q_e, q_g} U = v\left((q_e + q_g)m\right) + \theta_e q_e + \theta_g q_g + y - p_e q_e - p_g q_g.$$

Because the two fuels are perfect substitutes, motorists choose either to fuel with E10 or E85, but not both. A motorist chooses to fuel with E85 if the net utility benefit per gallon of E85 is greater than the net utility benefit per gallon of E10:

$$\theta_e - p_e \geq \theta_g - p_g.$$

Following the notation of Anderson (2012), we let $p \equiv p_e - p_g$ be the E85 price premium (or discount if negative) and $\theta \equiv \theta_e - \theta_g$ be the motorist's willingness to pay (or the amount to compensate the motorist if negative) for E85 as a substitute for E10. Thus, we can restate the decision of motorists to choose E85 if their willingness to pay for E85 over E10 exceeds the price premium they face at the pump, i.e., $\theta \geq p$.

Even though a motorist makes her fuel choice based on the difference in prices and her own preference parameter θ, the quantity of fuel demanded and in turn the motorist's miles driven depend only on the price of the fuel chosen. The first order conditions of the utility maximization problem show that, conditional on the motorist choosing fuel $j \in \{e, g\}$,

$$v'\left(q_j^* \cdot m\right) \cdot m + \theta_j - p_j = 0$$

To obtain the motorist's choice of miles driven and fuel demand, we re-write the equation:

$$v'\left(q_j^* \cdot m\right) = \frac{p_{j-}\theta_j}{m}$$

The motorist's choice of miles driven is $M^* = q_j{}^* \cdot m$. Solving the above equation for M^* yields:

$$M^* = v'^{-1}\left(\frac{p_{j-}\theta_j}{m}\right),$$

and the motorist's demand for fuel type j , is $q_j{}^* \equiv M^*/m$.

Aggregate Demand

To formally aggregate individual behavior and set up the empirical section, a few more assumptions are employed. The model assumes that each E85 station serves its own market of flex motorists, meaning that each E85 station is a local monopolist for E85, and the price of E85 at other stations does not affect the station's market size. Motorists in the station's market are aware of the prevailing E85 and E10 prices, and if they choose to fuel with E85, they visit the E85 station. If they choose E10, they may visit the E85 station (all E85 stations in Minnesota supply E10) or they may choose a nearby E10 (only) station. Note that an FFV motorist may be within the market of an E85 station even if there is an E10 station more directly along the motorist's normal driving path. As long as the motorist is aware of the E85 station, and the E85 station is not too far off of the motorist's normal driving path, then the motorist is within the station's market, and if the E85 premium is low enough, the motorist will visit the E85 station and choose E85 (Houde 2012).

The model assumes that motorist demand for miles is perfectly inelastic in the short run, and, without loss of generality, that motorists are heterogeneous with fuel demand q and willingness to pay for ethanol θ jointly distributed among motorists according to the joint probability density function (pdf) given by $f(q,\theta)$. The total quantity of E85 demanded from an E85 fuel station can be calculated as 1) the number of FFV motorists in the station's market multiplied by 2) the average fuel consumption among those motorists that choose ethanol multiplied by 3) the fraction of those motorists whose willingness to pay for ethanol exceeds the station's E85 price premium. Algebraically, this can be written as

$$Q = N \int_p^\infty [\int q f(q,\theta) dq] d\theta = N \int_p^\infty \boldsymbol{E}\,(q|\theta) f(\theta) d\theta,$$

where N is the number of flex motorists in the station's market, $\boldsymbol{E}(q|\theta) \equiv \int qf(q|\theta)dq$ is the expected fuel demand conditional on willingness to pay θ, and the expression is simplified using the fact that the joint pdf is the product of the conditional and marginal probability densities: $f(q,\theta) \equiv f(q|\theta) \cdot f(\theta)$. By multiplying and dividing by the unconditional expected fuel demand $\boldsymbol{E}(q)$ the expression can be further simplified:

$$Q = N \cdot \boldsymbol{E}(q) \cdot \int_p^\infty \frac{\boldsymbol{E}(q|\theta)}{\boldsymbol{E}(q)} f(\theta)d\theta = N \cdot \boldsymbol{E}(q) \cdot \int_p^\infty h(\theta)d\theta,$$

where $h(\theta) \equiv \boldsymbol{E}(q|\theta)/\boldsymbol{E}(q) \cdot f(\theta)$.

Anderson (2012) notes that $h(\theta) \geq 0$ and that $h(\theta)$ integrates to one, making it a proper pdf itself. One can think of $h(\theta)$ as the marginal pdf of willingness to pay for E85 among flex motorists, but instead of using the joint distribution with fuel demand given by $f(q,\theta)$, the distribution $h(\theta)$ puts weights on motorists according to fuel consumption. Defining $H(\theta)$ as the cumulative distribution function (cdf) associated with the pdf $h(\theta)$ allows us to rewrite aggregate ethanol demand:

$$Q = N \cdot \boldsymbol{E}(q) \cdot \int_p^\infty h(\theta)d\theta = N \cdot \boldsymbol{E}(q) \cdot [1 - H(p)]. \quad (2)$$

The model provides a direct mapping from the cdf of willingness to pay for E85 among flex motorists (weighted by volume of fuel demanded) to the station-level demand for E85. Taking the natural log of both sides yields a linear expression that provides the basis for the estimating equation we discuss in the next section:

$$\ln Q = \ln N + \ln \boldsymbol{E}(q) + \ln\big(1 - H(p)\big). \quad (3)$$

Empirical Model

For expositional purposes, we begin this section by re-writing theoretical equation (3) as:

$$\ln Q_{e_{it}} = \ln N_{it} + \ln \boldsymbol{E}(q_{it}) + \ln[1 - H(p_{it})]. \quad (4)$$

$Q_{e_{it}}$ is the quantity of E85 (in E10 energy-equivalent gallons) sold by E85 station i in month t, the product $N_{it} \cdot \boldsymbol{E}(q_{it})$ represents the total demand for E10 and E85 by flex motorists in the market of station i in month t, and $[1 - H(p_{it})]$ is the share of flex motorists (weighted by volume of fuel demanded) in the market of station i in month t who choose E85, given the E85 price premium p_{it}.

We assume that the volume-weighted distribution of willingness to pay for E85 is the same for all E85 station markets, remains constant over time, and follows a logistic distribution with mean μ and variance σ^2. This is unlike Anderson (2012) who assumes an exponential distribution of willingness to pay, focuses on the right tail of the distribution, and is unable to make out-of-sample predictions for lower E85 premiums.

The share of flex motorists who choose E85 at a given station in a given month is a function of only the station's monthly E85 premium. The logistic distribution has a sensible shape; it is symmetric, unimodal, and its support is all real numbers. Compared to the normal distribution, the logistic distribution has more mass on its tails, which is consistent with previous evidence of a large dispersion of willingness to pay for E85, and the cdf can be written in closed form. Letting $s = \sqrt{3}\sigma/\pi$, the cdf of the logistic distribution is

$$H(p; \mu, s) = \frac{1}{1+\exp\left(-\frac{p-\mu}{s}\right)},$$

where recall that p is the E85 premium.

Next we focus on the total demand for E10 and E85 by flex motorists given by $N_{it} \cdot \boldsymbol{E}(q_{it})$. The number of FFVs in a given station's market in a given month and the mean fuel demand of those vehicles are not observable. We therefore rely on a set of observable variables that explain the total fuel demand by FFVs in the E85 market of station i in month t. Specifically, we express the log of total fuel demand by flex motorists as

$$\ln N_{it} + \ln \boldsymbol{E}(q_{it}) = \gamma' \boldsymbol{X}_{it} + \delta_i + \varepsilon_t + \zeta_t + \omega_i \cdot t,$$

where

$$\gamma' \boldsymbol{X}_{it} \equiv \gamma_1 \ln(\#E85stations)_{it} + \gamma_2 M1_{it} + \gamma_3 M2_{it} + \gamma_4 M3_{it} + \gamma_5 M4_{it}.$$

$\ln(\#E85stations)_{it}$ is the log of the total number of E85 stations operating in the same county as station i in month t, $M1_{it}$, $M2_{it}$, $M3_{it}$, and $M4_{it}$ are dummy variables for the first four months that a station sells E85, δ_i is a station fixed effect, ε_t is a month fixed effect, ζ_t is a year fixed effect, and $\omega_i \cdot t$ is a station-specific time trend. We use these measures to estimate the size of the market for each E85 station because we do not have monthly, time-series, local-level data for the number of E10 stations, the number of flex motorists, or other relevant population characteristics. We rely on the station fixed effect and the station-specific time trend variables to capture these and other attributes of the station and surrounding market for fuel.

The station fixed effects control for unobserved station characteristics that remain constant over time. These may include the presence of E85 signage, the prominence and convenience of the station's E85 pump(s), the station's location (distance to a major highway, big city or small town, etc.), and possibly other demographic characteristics that are potential determinants of local demand such as infrastructures or the availability of public transport. The month fixed effects control for seasonality in motor fuel consumption, and the year effects control for longer-term, market-wide variation in motor fuel consumption, such as the decrease in fuel consumption observed during the last recession. Finally, the station-specific time trends control for effects correlated with time such as growth in the local stock of FFVs or a gradual increase in the local median income.

The model does not control for fuel prices at nearby E85 stations. This is reasonable if E85 search costs are relatively high for consumers and/or E85 stations are relatively spread out. However if there is more than one E85 station in a relatively small area, and prices are displayed prominently, motorists may choose to forego their usual E85 station and choose a neighboring E85 station instead, and this would be problematic for our model. Fortunately, most E85 stations in Minnesota are relatively far from one another, and both E85 and E10 prices are very similar day-to-day among nearby stations, so in general there is not much to be gained by motorists from searching for the station with the lowest fuel prices. Even in cases where fuel stations offering E85 are near one another, the gains that flex motorists can expect from searching are not likely to last long because fuel stations quickly respond to competitors' prices.

The estimating equation is

$$\ln Q_{e_{it}} = \beta_0 + \gamma' \boldsymbol{X}_{it} + \delta_i + \varepsilon_t + \zeta_t + \omega_i \cdot t + \ln\left[1 - \frac{1}{1+exp\left(-\frac{p_{it}-\mu}{s}\right)}\right] + u_{it}, \tag{5}$$

where u_{it} is the residual; and β_0, the γ-vector, δ_i, ε_t, ζ_t, ω_i, μ, and s are coefficients to be estimated. The model in (5) is similar to the empirical model estimated by Anderson (2012).

Extensions of the Model

We perform several robustness checks and extensions on this model. First, the model in (5) assumes that the size of the market for an E85 station is not affected by the station's E85 premium. That is, the model assumes that motorists do not go out of their way to seek out E85 stations when the E85 premium is particularly favorable. Recall that the size of an E85 station's market is the total fuel demand by the flex motorists in the area. The empirical model in (5) explains a station's market size with location, signage, brand, and other factors captured by the station fixed effects and other controls in the model, but omits fuel prices.

This assumption potentially misses an important characteristic of retail fuel markets: motorists are not stationary when they are consuming fuel, as pointed out in Houde (2008). Motorists encounter many retail fuel stations along their normal driving route, and may choose one that is further out of their way if the price is favorable. If the size of a station's market depends on its E85 premium, then omitting the premium term from that part of the model will bias estimates of the distribution of willingness to pay. In the first extension of the model, we explore the robustness of our basic results to the inclusion of the E85 premium to explain the size of an E85 station's market. In this version of the model, flex motorists may drive out of their way to purchase E85 in months when the E85 premium is low. We assume that the marginal benefit of the money saved is decreasing, and we model the effect of the E85 premium on the size of the market as being linear in logs.

Second, the model in (5) assumes a perfectly inelastic demand in the short run and as such does not include the price of E85 or E10 to explain consumption volumes. If false, this assumption could potentially bias our results. In particular, if consumption volumes are sensitive to fuel prices in the short run and fuel prices are correlated with the E85 premium, then the zero short-run elasticity assumption would bias our estimate of the distribution of willingness to pay. We explore the impact of the short-run assumption in an alternative specification of the econometric model where we allow the absolute fuel price to affect fuel consumption.

Estimating the Parameters of the Willingness to Pay Distribution

We could obtain estimates of μ and s directly using a nonlinear estimator. But given the size of our data sample and the number of fixed effects in our model, estimation of the model's parameters becomes computationally intensive and the results are sensitive to the choice of starting values. We instead use a linear specification of the empirical model which, in addition to making numerical convergence easier, allows us to deal with potentially endogenous E85 prices more conveniently.

To linearize the empirical equation, we use a second-degree Taylor series approximation of $\ln[1 - H(p_{it})]$. It is reasonable to assume a mean willingness to pay for E85 that is not too far from zero, where the cost per mile is the same for both fuels. If on average motorists' valuation of E85 relative to E10 deviates from the parity price, we do not expect it to deviate by much because E10 and E85 are overall very similar products and attributes that differentiate them likely represent only a small share of the average motorists' valuation. Taking a second-degree Taylor approximation of $\ln[1 - H(p_{it})]$ around zero yields

$$\ln[1 - H(p_{it})] \approx \ln[1 - H(0)] - \frac{H'(0)}{1-H(0)} \cdot p_{it} + \left(\frac{H'(0)^2}{\left(1-H(0)\right)^2} - \frac{H''(0)}{1-H(0)}\right) \cdot \frac{{p_{it}}^2}{2}. \tag{6}$$

Writing the linear and the quadratic terms of the Taylor approximation as β_1 and β_2, the linearized version of equation (5) is

$$\ln Q_{e_{it}} = \widetilde{\beta_0} + \beta_1 p_{it} + \beta_2 {p_{it}}^2 + \gamma' \boldsymbol{X}_{it} + \delta_i + \varepsilon_t + \zeta_t + \omega_i \cdot t + u_{it}, \tag{7}$$

where $\widetilde{\beta_0} = \beta_0 + \ln[1 - H(0)]$. β_1 and β_2 are parameters to be estimated that are functions of the parameters μ and s of the distribution function $H(p_{it})$. More specifically, given that we use a logistic distribution function, the expressions for the parameters β_1 and β_2 are

$$\beta_1 = \frac{-1}{s(1 + e^{\mu/s})}; \tag{8}$$

$$\beta_2 = \frac{-e^{\mu/s}}{2s^2(1 + e^{\mu/s})^2}. \tag{9}$$

Solving (8) and (9) allows us to obtain estimates of μ and s (and in turn σ) from linear estimation.

In the first extension of the empirical model, we use a third-order approximation of $\ln[1 - H(p_{it})]$, and we use the coefficients on the E85 premium squared and E85 premium cubed to recover estimates of μ and σ. We do this to allow the E85 premium to linearly affect the log of the size of an E85 fuel station's market. That is, if a motorists' decision to enter a particular E85 fuel station's market is a function of the E85 premium that is relatively linear in logs, then the coefficient β_1 captures both the decision of motorists to enter the E85 station's market and the decision of motorists already in the station's market to choose E85 instead of E10. Under this assumption, the coefficients for the E85 premium squared and cubed solely capture willingness to pay. With a third-degree Taylor approximation and a logistic distribution function for willingness to pay, the expression for β_3 is

$$\beta_3 = \frac{(1 - e^{\mu/s})e^{\mu/s}}{6s^3(1 + e^{\mu/s})^3} \tag{10}$$

In all of the extensions we perform using the cubic model, we use estimates of β_2 and β_3, and we solve equations (9) and (10) numerically to find values for μ, s, and in turn σ.

Identification and Estimation

We estimate the econometric model using both ordinary least squares (OLS) and two-stage least squares (2SLS). In the 2SLS estimation, we instrument for the E85 premium, as well as the log E85 price when it is included in the model, to address the potential endogeneity problem. Our 2SLS estimation approach uses supply-side variables to identify the parameters of the distribution of willingness to pay by flex motorists.

We perform OLS estimation as well because it is possible that the estimates for μ and s are not severely biased. Stations often set E85 fuel prices based on the wholesale E85 price, diminishing the effect of local E85 demand shifts correlating with station-level E85 prices and premiums. However there is a potential that some station-level E85 demand shocks are correlated with the station's E85 premium, so we prefer to estimate the model using 2SLS.

Another reason to prefer 2SLS is to correct for endogenous measurement errors in the E85 premiums. While we observe each station's E85 price, we do not observe the station's local market E10 price. Instead, we rely on the statewide monthly average E10 price to calculate a station's E85 premium. The measurement error is the difference between the actual local E10 price and the statewide average E10 price. If local E10 prices are correlated with local E85 prices, then the measurement errors are correlated with the E85 premiums. This means our OLS estimates could suffer from attenuation bias, but our 2SLS estimates do not. For example, if the price of E10 and E85 in some local market are both high in a given month, and the local E10 price is higher than the statewide average E10 price, then the observed E85 premium is higher than the actual E85 premium, and estimates overstate the share of motorists who choose E85 when the premium is high. Alternatively, if the local E10 and E85 prices are low in some month in some market such that the local E10 price is lower than the statewide average E10 price, then the observed E85 premium would be less than the actual premium, and estimates would understate the share of motorists who choose E85 when the premium is low. Therefore our OLS estimates of the distribution of willingness to pay for E85 could be biased to show a higher variance in willingness to pay for E85

To instrument for potentially endogenous or mismeasured E85 premiums, we begin with a set of simple instruments we believe to be uncorrelated with local, short-run demand shifts, but correlated with the station's E85 premium. To instrument for a station's E85 premium, based on Anderson (2012), we use the wholesale price of E10, and the wholesale price of E85[2], and we interact these two price series with the number of E85 stations per square mile and the number of all fuel stations per square mile in the same county as the station. These interactions create four variables that capture not only how wholesale prices affect retail prices, but also how local competition affects how retailers respond to those wholesale prices. A retailer in an area with many E85 stations may need to lower her E85 price when the wholesale price drops whereas an E85 retailer who faces less competition may be able to keep her E85 price high. In addition to these four (4) instruments, we include (5) the wholesale price of corn,

[2] Wholesale E85 prices are calculated as weighted averages of the wholesale (refiner) E10 price and the wholesale (rack) ethanol price minus the value of the RIN:

$$wholesale\ E85\ price = \alpha * wholesale\ E10\ price + (1-\alpha) * (wholesale\ ethanol\ price - RIN\ price).$$

The weights are according to the E85 Handbook's nominal ethanol content of E85 in Minnesota for a given month:

$$E85\ ethanol\ percent = \alpha * 0.10 + (1-\alpha).$$

(6) a one-month lag of the log of the station's E85 price, (7) a one-month lag of the log of the station's E85 quantity sold, and (8) a one-month lag of the station's E85 premium.

Next, we use a more complex set of instruments. We generate these instruments in the same manner as Anderson (2012). In addition to the list of instruments described in the previous paragraph, we use the interaction of the wholesale E10 and E85 fuel prices with the station's brand and distance to supplier. Unfortunately, the more complex set of instruments comes at a cost as we do not observe brand or exact geographic location for all fuel stations, thus forcing us to remove observations where station-specific data are not available. We discuss the instruments and estimation sample further in the next section.

Data

We use monthly data from a variety of sources to estimate equation (7) for a large sample of E85 stations in Minnesota. The sample we use for our initial estimation consists of 15,235 monthly observations from 288 stations. Table 1 contains summary statistics of these data. All fuel prices and quantities are in E10-equivalent energy units, and all prices are in 2014 dollars. In this section, we explain the sources and calculations used to generate these data, and we also discuss the properties of the data.

Dependent and Independent Variables

The data for E85 prices and sales volumes come from MN DoC (2014). Each month, MN DoC surveys retail E85 stations all over the state. The stations report E85 sales volumes and revenues which are used by MN DoC to calculate volume-weighted monthly average prices. The complete dataset generously provided by MN DoC contains 21,357 observations from 413 E85 stations. The dataset includes station-specific variables for the monthly E85 sales volume, the monthly E85 price, and limited information about the location of the E85 fuel stations.

The data cover the period from October 1997 to August 2014. Not all stations report in every month, but the E85 stations that received government funding to pay for their infrastructure costs can have those loans partially forgiven by reporting, and many stations participate voluntarily. MN DoC also provide the total number of E85 stations operating in the state each month. The number of E85 stations in Minnesota grew from fewer than 10 to around 350 during the timespan of the data. On average about 54 percent of stations reported sales volumes and prices to MN DoC, as shown in Figure 2.

We use the data from MN DoC to tabulate the number of E85 stations in each county that respond to survey in each month, and we divide that number by the fraction of the statewide E85 stations that participate in the survey that month. This variable acts as a proxy for the number of E85 stations operating in each county in each month under the assumption that the proportion of E85 stations that report is the same across counties. Next, we generate dummy variables for the first, second, third, and fourth month after a station begins reporting. We assume the first month a station reports to MN DoC is the first month that the station sold E85. We use these variables to explain the size of a particular E85 station's market. Flex motorists in the area may take some time to learn of the existence of the E85 station and to observe the E85 premium.

We convert the E85 prices and sales volumes into gasoline energy-equivalent units. Almost all regular gasoline in Minnesota is E10 and contains roughly 10 percent ethanol during any given month of the year, but the amount of ethanol in the E85 fuel blend depends on the season. In the winters, a higher concentration of gasoline is needed to ensure proper starting in cold conditions. According to the E85 handbook published by the US Department of Energy (DOE), E85 in Minnesota contained between 70 and 79 percent ethanol for most of the duration of the data collection period – 70 percent in the winter months and only reaching 79 percent in July (DOE 2008). Using these blend concentrations, and assuming that pure ethanol has two-thirds the energy content per volume as pure gasoline, we calculate conversion factors for each month ranging from 1.26 in January to 1.31 in July.

To calculate the E85 premium, we obtain monthly data on the retail price of regular unleaded E10 gasoline in Minnesota from the US Energy Information Administration (EIA). EIA surveys around 800 retail locations across the country each week to obtain price data, and it also uses monthly sales reports from petroleum resellers and retailers (EIA 2013a and EIA 2013b). These price data and the E85 price data from MN DoC include all taxes and are the end prices paid by the consumer. EIA combines these price data with other sales and population data to calculate weighted average price estimates at the state level.

We convert the retail E85 prices and the retail E10 prices into August 2014 dollars using monthly CPI data from the US Department of Labor Bureau of Labor Statistics (BLS 2014). Figure 3 shows the energy-adjusted real retail price of E85 from each station in each month in our sample along with the statewide average real retail price of E10.

We calculate the E85 premium as the difference between the energy-equivalent real retail price of E85 and the real retail price of E10. Figure 4 shows the E85 premiums at the E85 stations in our sample. Each individual dot in Figure 4 shows the E85 premium at one station in one month, and the line shows the average E85 premium from among the reporting stations. When the E85 premium is positive, the energy-adjusted price of E85 is higher than the price of E10. From October 1997 through August 2014, only in March 2014 the average E85 premium was negative in energy-equivalent terms. The average E85 price had fallen to within just two cents of parity with E10 in November, 2007. But by the end of 2012, E85 was on average sixty cents per gallon more expensive than E10 because of high corn prices. Corn and ethanol prices fell in 2013 and 2014, and the E85 premium fell sharply as well. We also note that although the statewide average energy-adjusted E85 premium has almost always been higher than the statewide average price for E10, there are thousands of instances where individual stations have offered E85 at a discount relative to E10 in a certain month.

Instrumental Variables

As mentioned briefly in the previous section, our initial set of simple instruments consists of the wholesale prices of E10, E85, corn, and the density of E85 and all (E10) fuel stations in the same county, as well as one-month lags of the log of the station's E85 quantity sold, the log of the station's E85 price, and the station's E85 premium. EIA provides monthly data on the wholesale price of E10 in Minnesota (EIA 2014c).

Monthly data for the wholesale price of ethanol are obtained from the Nebraska Energy Office (NEO). NEO reports ethanol average rack prices in Omaha, NE each month. The rack price is the price for truck quantities of pure ethanol charged by ethanol producers to blenders, resellers, and other various clients at the given location (NEO 2014). Because Omaha is relatively close to Minnesota, the Omaha price is likely close to the price paid in Minnesota. We subtract the monthly average RIN price from the wholesale ethanol price. We obtain RIN prices from the Oil Price Information Service (OPIS). We calculate the wholesale price of E85 in each month as the weighted average of the wholesale E10 price and the wholesale (rack) price of ethanol minus the RIN value. The weights are based on the monthly average ethanol concentration in E85 reported by DOE (2008). We then convert the wholesale E85 price series into E10 energy-equivalent dollars and convert both the E10 and the E85 wholesale fuel price series into August 2014 dollars. We obtain wholesale corn prices from the Chicago Board of Trade (CBOT) by taking the monthly average of corn futures prices.

As in Anderson (2012), we interact the wholesale E10 and E85 price series with measures of local competition. We calculate the density of E85 stations and the density of all fuel stations in the county where the station is located. We obtain the number of E85 stations in each county from a list maintained by the Alternative Fuels Data Center (AFDC) that provides a snapshot of the E85 retail stations operating in Minnesota in September, 2013. The number of E10 retail fuel stations in each county is obtained from MN DoC in a separate dataset, and also represents a snapshot of the operating stations in Minnesota in September, 2013. We obtain the area (in square miles) of each county in Minnesota from the US Census, and we calculate the E85 and E10 station densities as the number of stations per square mile. The intuition for using these variables as instruments is that retailers facing stiff competition may be more inclined to behave as competitive firms who set their price equal to the marginal cost. On the other hand, E85 retail stations not facing such competition may behave as local monopolists, and their retail prices may therefore be less tied to the wholesale prices.

The original dataset from MN DoC provides the county where each station is located but not the exact geographic location. However, the AFDC's list of E85 stations provides the stations' exact geographic coordinates, the stations' names, the station's county, and the date the station first started selling E85. By cross-referencing the AFDC list of stations with the data from MN DoC, we are able to infer which E85 price/quantity series belong to which E85 stations based on the station's county and the month and year the station began selling E85. Using this method, we are able to positively identify 306 of the 413 stations in the original dataset. The remaining stations could not be identified for one of two reasons. First, we were not able to identify stations that closed before September, 2013 and thus were not on the AFDC's list of E85 stations. Second, we were not able to uniquely identify stations from the same county with the same start date (month and year). For reasons we discuss in the next section, we limit the initial estimation sample to 288 of the 413 stations in the dataset. We are able to positively identify 246 of those 288 stations.

For the identified stations, we measure an individual E85 retailer's supplier-relationship by calculating the log of the distance (in miles) from the station to the nearest ethanol blending terminal. In addition to capturing a supplier-relation effect, this distance variable also captures the direct, supply-side, transportation cost of supplying the fuel to the station. We create dummy variables for each brand affiliation. Any brand with at least two stations has its own dummy variable, and any station with a unique brand or a brand we could not identify we designated as, 'Other'. This method generates 16 brand categories for the 246 stations.

To construct the more complex set of price instruments that captures how individual retail stations respond to changes in supply-side costs, we interact the wholesale E85 prices and wholesale E10 prices with 1) the number of E85 stations per square mile in the county, 2) the number of E10 stations per square mile in the county, 3) the logged distance in miles to the nearest blending terminal, and 4) the 16 brand dummies. These interactions produce a total of 38 instrumental variables. We also keep the rest of the initial instruments: the wholesale price of corn, and the one-month lag values of the station's log of E85 quantity sold, log of E85 price, and E85 premium. The instruments allow us to remedy the endogeneity problem by modeling retail E85 pricing behavior that is exogenous to local, short-run shifts in E85 demand. In addition, instrumenting for the E85 premium in this fashion allows us to correct for the potential measurement errors in the premium discussed in section IV.

Estimation Sample

Although the original survey data contain more than 21,000 monthly observations from 413 stations, our initial estimation sample consists of 15,235 observations from 288 stations. The reasons for dropping observations are that 1) we remove any E85 price or quantity observations that are extreme outliers likely resulting from reporting error (such as months where the total quantity sold or average price is zero), 2) we use one-month lagged values as instruments so any observation without an observation the preceding month is incomplete, 3) we only use observations from stations with at least ten complete observations, and 4) we only use observations from the most recent eight years of data – from September 2006 to August 2014.

We do not include observations from between 1997 and 2005 because almost all of the E85 sales during that period were to government vehicles required by law to use E85. Neither FFVs nor E85 infrastructure were common during that period and data from that time likely misrepresent the preferences of today's FFV motorists. To examine how our choice of start date affects our estimates, we also estimate model using estimation samples with observations starting in 2004, 2008 and 2010. Finally after removing observations from stations we cannot identify, the dataset contains 13,941 observations from 246 identified stations.

Table 1 shows summary statistics for the initial estimation sample of 15,235 observations from all of the E85 stations. Table 2 shows the same summary statistics for the 13,941 observations from the identified stations. Comparing Table 1 to Table 2, the summary statistics do not suggest that the data samples are decidedly different from each other. However, to examine the possibility of sample selection and to see how it affects our results, we estimate the model with OLS and with 2SLS using the simple instruments using both the full dataset containing observations from all stations as well as the subset of the data containing observations from only identified stations.

Econometric Estimation and Results

We estimate the model in (7) under several specifications to verify the robustness of our estimates to the choice of instruments, the estimation sample selected, and the assumptions about motorists' motives to fill at a fuel station. What is common in all specifications is that we apply the standard one-way fixed effects model by subtracting the each stations' mean observations and performing OLS on the transformed data (Baltagi 2013). We choose a fixed effect model over a random effect model because of the potential correlation between a station's fixed-effect and its premium. We do not estimate the model in first difference because with our unbalanced dataset it would cause the loss of a large number of observations.

We estimate the model using either all 15,235 observations or the 13,941 observations for which we have brand and location information. Both data samples cover the period between September 2006 and August 2014. We label the sample with 288 stations as "All" and the sample with 246 stations as "Identified". In Model 1, we estimate the model using OLS and the sample with All stations. In Model 2, we estimate the model using OLS and the sample with Identified stations. Then in Models 3 and 4, we estimate the model using 2SLS and the simple set of instruments with the All and Identified samples respectively. In Model 5 we use the complex set of instruments with the Identified sample.

Table 3 shows the results. The table shows coefficient estimates for the E85 premium, the E85 premium squared, and the log of the number of E85 stations operating in the same county. Standard errors are in parentheses. The table also contains the estimates of the means and standard deviations of the distribution of willingness to pay implied by the coefficients for the premium and the premium squared. Values of μ and σ are calculated solving equations (8) and (9) and their standards errors are calculated using the delta method. Recall that we include fixed effects for each station, year effects, month effects, station-specific time trends, and dummy variables for the first four months the station sells E85. We do not report the coefficient estimates for these variables for each of our estimations, but we note that E85 demand is highest in the months of May, June, July, and August and lowest in December, January, and February. The year effects are the most negative (compared to 2006) in 2009, 2010, 2013, and 2014. Appendix A contains complete tables of results.

In every model we estimate, the coefficients for the premium and the premium squared are both negative. This means that at the average premium observed in the dataset, the marginal distribution of willingness to pay is declining with respect to the premium. This is consistent with the observation that E85 has been priced at a premium over E10 that is larger than the mode of the distribution of willingness to pay. Given that in the case of the logistic distribution function that the mode equals the mean and the median, the implication is that the premium on E85 has been on average larger than the mean willingness to pay for E85 over E10. Another common result through the models is that the coefficient estimate on the log of the number of E85 stations in the county is negative, and it ranges between -0.042 and -0.057 depending on the estimation sample and model specification. This implies that increasing the number of E85 stations in a county by 10 percent will reduce the E85 sales volumes of the other E85 stations in the county by about 0.5 percent, conditional on the location decisions of the new E85 stations. The relatively low estimated decrease in sales volumes from additional E85 stations suggests that E85 markets are relatively distinct or that E85 retailers choose to locate in areas where E85 is not already available.

Models 1 and 2 use OLS and thus treat the E85 premium as exogenous. The results of the OLS models suggest that the average FFV motorist requires a discount on E85 of about \$1.45 or \$1.28 per gallon depending on the estimation sample. The distribution of willingness to pay is relatively wide with a standard deviation of about \$1.67 or \$1.64 per gallon.

As mentioned earlier, there are good reasons, including measurement errors, to suspect that the premium is endogenous to the volumes of E85 sold. So we estimate the model using 2SLS and our set of basic instruments which do not require us to identify the individual fuel stations in Models 3 and 4. In the case of the All sample, the estimate of μ rises from -\$1.45 per gallon to -\$1.21 per gallon, but in the case of the Identified sample, the OLS estimate and the 2SLS estimate of μ are the same -\$1.28 per gallon. More notable is the difference between the OLS and 2SLS estimates of σ, the standard deviation of the willingness to pay distribution. When we use 2SLS, the estimates of σ fall to \$1.22 per gallon and \$1.27 per gallon, about 25 percent less than the OLS estimates. This is what we would expect given our discussion of the potential measurement errors. In Model 5, when we expand to the complex set of instruments that require identification of the fuel stations, the estimate of σ remains relatively unchanged at \$1.23 per gallon, but the estimate of μ increases to -\$1.00. Appendix B shows summary results for the first stage regressions of our 2SLS models.

Next we extend the model to allow the possibility that a station's E85 premium influences the size of the station's market. We suppose that if the station's E85 premium is low enough, then motorists who would not normally consider visiting the station might choose to visit the station and choose to fuel with E85 simultaneously. For example a flex motorist may be willing to purchase E85 when it is offered at a premium of, say, \$0.20 if it were offered nearby, but the motorist normally visits fuel stations that sell only E10, and must incur some additional cost to visit the station that sells E85. Then we could imagine that if the E85 premium fell sufficiently, the flex motorist might decide to visit the E85 station.

However the E85 premium required to induce the motorist to travel to the E85 station and fuel with E85 would be lower than the motorist's true willingness to pay for E85. In this way, the station's E85 premium affects the share of motorists who choose E85 among its regular flex motorist patrons, and it also affects the size of the E85 station's market.

To capture this behavior in our model, we assume that the linear premium term affects both the size of the station's E85 market as well as the share of flex motorists within the market who choose E85. We include a term in the regression for the E85 premium cubed, and we use the second and third degree premium coefficients to recover estimates for the mean and variance of the willingness to pay distribution. Note that we assume the coefficients on the premium squared and cubed are not biased, and this implicitly assumes that the effect of an E85 station's premium on the size of its market is linear in logs, consistent with decreasing marginal utility of money.

We compare the estimates from the empirical specification with only the premium and premium squared to the estimates obtained from the third-degree Taylor-series approximation model in Table 4. We include the results of Model 2 and Model 5, the models using the squared version of the estimating equation estimated with OLS and 2SLS with the Identified sample. Then we estimate the cubic version of the model using OLS and 2SLS, and we use the second and third degree E85 premium coefficient estimates to obtain estimates of μ and σ. Just as was the case with the squared premium model, the 2SLS coefficient estimate of μ is higher than the OLS estimate – from -\$0.65 with OLS to -\$0.57 with 2SLS. In the cubic model, the estimates of σ are relatively similar for both OLS and 2SLS estimation: \$0.47 and \$0.49. The estimates of μ in the cubic model are markedly higher than in the squared model, and the estimates of σ are markedly lower. We calculate the Bayesian Information Criterion (BIC) for the squared and cubic models estimated with both OLS and 2SLS. In both the OLS and 2SLS case, the cubic model is preferred to the squared model, indicating that the cubic model provides a significantly better fit even after adjusting for the addition of a regressor. For this reason and because we believe the cubic extension of the model more accurately represents reality, Model 7 is our preferred model which we use as a baseline to perform the other extensions of the model. We use the calculated parameter values for μ and σ from Model 7 to plot the cdf of willingness to pay for E85 as a substitute for E10 among flex motorists, weighted by fuel consumption, shown in Figure 5. When E85 and E10 are equivalently priced in energy-equivalent dollars, about 11 percent of flex motorists choose E85.

The first extension we make to the baseline Model 7 is that we relax the assumption that fuel demand is perfectly inelastic in the short run. We begin by imposing sensible values for the elasticity of

E85 fuel other than zero, and we also estimate the price elasticity of demand directly in the log-log model. Table 5 shows the results. We find that the greater the magnitude of the elasticity parameter we impose on the log of E85 price, the smaller the magnitude of the coefficient on the E85 premium and E85 premium cubed, but the larger the magnitude of the coefficient on the E85 premium squared. This leads to marginally higher estimates for the mean willingness to pay – from -$0.56 when the short-run elasticity of demand for E85 is fixed at -0.10 up to -$0.52 when the elasticity is fixed at -0.5. The change in coefficient estimates also results in a slightly higher estimate for the standard deviation – from $0.50 to $0.53. We also estimate the short-run elasticity of demand freely using 2SLS, instrumenting for the log of the E85 price along with the E85 premium using the same set of complex instruments we use in our baseline model. The estimate for the coefficient on log E85 price is 0.846 which is troubling because we expect the sign to be negative. This could be an indication of weak instruments or model misspecification. In this empirical specification, the estimated mean willingness to pay for E85 changes to -$0.62 per gallon with standard deviation $0.41 per gallon.

We also estimate the model using different starting dates. Recall that our baseline estimates use eight years of data – from September 2006 to August 2014. We do not make this choice arbitrarily, but rather because we suppose that this range of the data best represent the preferences of U.S. flex motorists. As described in section II, before 2006, there were few E85 stations, the sales at those stations were limited, few private motorists owned FFVs, and most E85 sales were to government fleet vehicles. To investigate the effect our choice of start date has on our estimates, we also estimate the baseline empirical specification using datasets with different starting years. Specifically, along with our original estimation sample starting in 2006, we use samples starting in 2004, 2008 and 2010. Table 6 shows the results. The sample starting in September 2010 has 8,082 observations from 222 stations. To identify whether or not the estimates are affected by the different start dates and not just the addition or subtraction of stations that may have closed or stopped selling E85, we only use these 222 stations for the estimation sample for all of the start dates. So for example the sample dataset dating back to September, 2004 has 14,643 observations from those same 222 stations. The coefficient and parameter estimates are slightly different between the estimation samples. The mean willingness to pay for E85 ranges from -$0.53 per gallon for the sample starting in 2004 to -$0.59 per gallon for the sample starting in 2010. The estimated standard deviations range from $0.42 per gallon to $0.47 per gallon. The estimates differ only slightly between estimation samples, and all of the estimates are within one standard error of the estimates from our baseline model.

To summarize the results reported in this section, we find that the average flex motorist prefers E10 when the two fuels are priced at an energy-equivalent level, but that fuel-switching behavior spans a wide range of prices. Whether we use simple or complex price instruments has a small impact on our estimates, as does the choice to use only observations from identifiable stations instead of all observations. This can be seen in Table 3 by comparing the results of Models 3, 4, and 5. We also find that our choice of a start date for the estimation sample has a relatively small impact on our estimates, as shown in Table 6. The biggest impact on the results comes from using 2SLS estimation instead of OLS, and from using the cubic version of the model instead of the squared version. In our preferred model, we allow an E85 station's premium to affect its market size, and we include a term for the E85 premium cubed, applying a third-degree Taylor approximation of the nonlinear share function. When we estimate our preferred model with 2SLS and the complex set of instruments, we calculate mean willingness to pay for E85 $\mu = -\$0.57$ per gallon and $\sigma = \$0.49$ per gallon.

Conclusion

The demand for ethanol as a motor fuel is an important and debated topic in the United States in 2015. Only a few studies have attempted to estimate the demand for ethanol in the United States beyond the E10 blend wall, and those studies suffer from a lack of available data. Our model assumes flex motorists choose between E10 and E85 based on observed prices as well as personal preferences. We use a dataset of E85 prices and sales volumes from E85 fuel stations in Minnesota to estimate an empirical model of demand for E85 as a substitute for E10. We find that on average flex motorists discount E85 by about $0.57 per gallon relative to E10 when the price of E10 of $3.33 per gallon. In other words, if E85 were cheaper than E10 by about 17 percent (in E10 energy-equivalent terms), we would expect about half of flex motorists to choose E85. We also find that the motorists' fuel-switching behavior spans a wide range of values for the E85-E10 price difference, which we call the E85 premium. Our estimates indicate that 11 percent of flex motorists are willing to pay a positive premium to use E85 instead of E10, while 11 percent of motorists would not use E85 even if it were discounted by $1.14 per gallon.

Our estimates of the mean and variance of the distribution of consumer willingness to pay to use E85 as a substitute for E10 fall within the range reported in the related literature. However our study is the first to estimate the spread of preferences using station-level data, from the United States, from a time period when most sales were to private motorists, and when E85 was not significantly and persistently more expensive than traditional E10.

Our results suggest that ethanol quantities in excess of the blend wall could be consumed in the United States through E85 if (1) more motorists owned FFVs, (2) more retail fuel stations carried E85, and (3) E85 was priced more competitively with E10. Our calculations imply that even when E85 is priced slightly higher than E10 on a cost-per-mile basis, a sizable share of FFV motorists choose to fuel their vehicles with E85 given the option to do so.

References

Alternative Fuels Data Center. (2013). *Alternative fueling station locator.* Retrieved from http://www.afdc.energy.gov/locator/stations.html

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

Babcock, B. A., & Pouliot, S. (2013a). Price it and they will buy: How E85 can break the blend wall. *CARD Policy Brief 13, PB 11*.

Babcock, B. A., & Pouliot, S. (2013b). Impact of Sales Constraints and Entry on E85 Demand. *CARD Policy Brief 13, PB 12*.

Baltagi, B. (2013). *Econometric analysis of panel data.* Chichester, UK: John Wiley & Sons Ltd.

Bureau of Labor Statistics. (2014). *CPI Detailed Report, Data for August 2014, table 24.* Retrieved from http://www.bls.gov/cpi/cpid1408.pdf

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

Department of Energy. (2008). *Handbook for handling, storing, and dispensing E85.* Publication: DOE/GO-102008-2450, National Renewable Energy Laboratory.

Energy Information Administration. (2013). *Gasoline prices by formulation, grade, sales type*. Retrieved from http://www.eia.gov/dnav/pet/pet_pri_allmg_d_SMN_PTA_dpgal_m.htm

Energy Information Administration. (2014a). *Weekly retail gasoline and diesel prices*. Retrieved from http://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_smn_m.htm

Energy Information Administration. (2014b). *Refiner gasoline prices by grade and sales type*. Retrieved from http://www.eia.gov/dnav/pet/pet_pri_refmg_dcu_smn_m.htm

Energy Information Administration. (2014c). *Annual energy outlook 2014, table 58*. Retrieved from http://www.eia.gov/forecasts/aeo/tables_ref.cfm

Energy Information Administration. (2015). *Annual energy outlook 2015, table 12*. Retrieved from http://www.eia.gov/forecasts/aeo/tables_ref.cfm

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

Liu, C., & Greene, D. (2013). Modeling the demand for E85 in the United States. Available at: http://info.ornl.gov/sites/publications/Files/Pub45616.pdf

Minnesota Department of Commerce. (2014). *E85 fuel use data*. Retrieved from http://mn.gov/commerce/energy/topics/clean-energy/Vehicles-Fuels/Ethanol.jsp

Nebraska Energy Office. (2014). *Ethanol and unleaded gasoline average rack prices.* Retrieved from http://www.neo.ne.gov/statshtml/66.html

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

United States Census Bureau. (2014). *Minnesota state and county QuickFacts.* Retrieved from http://quickfacts.census.gov/qfd/states/27000.html

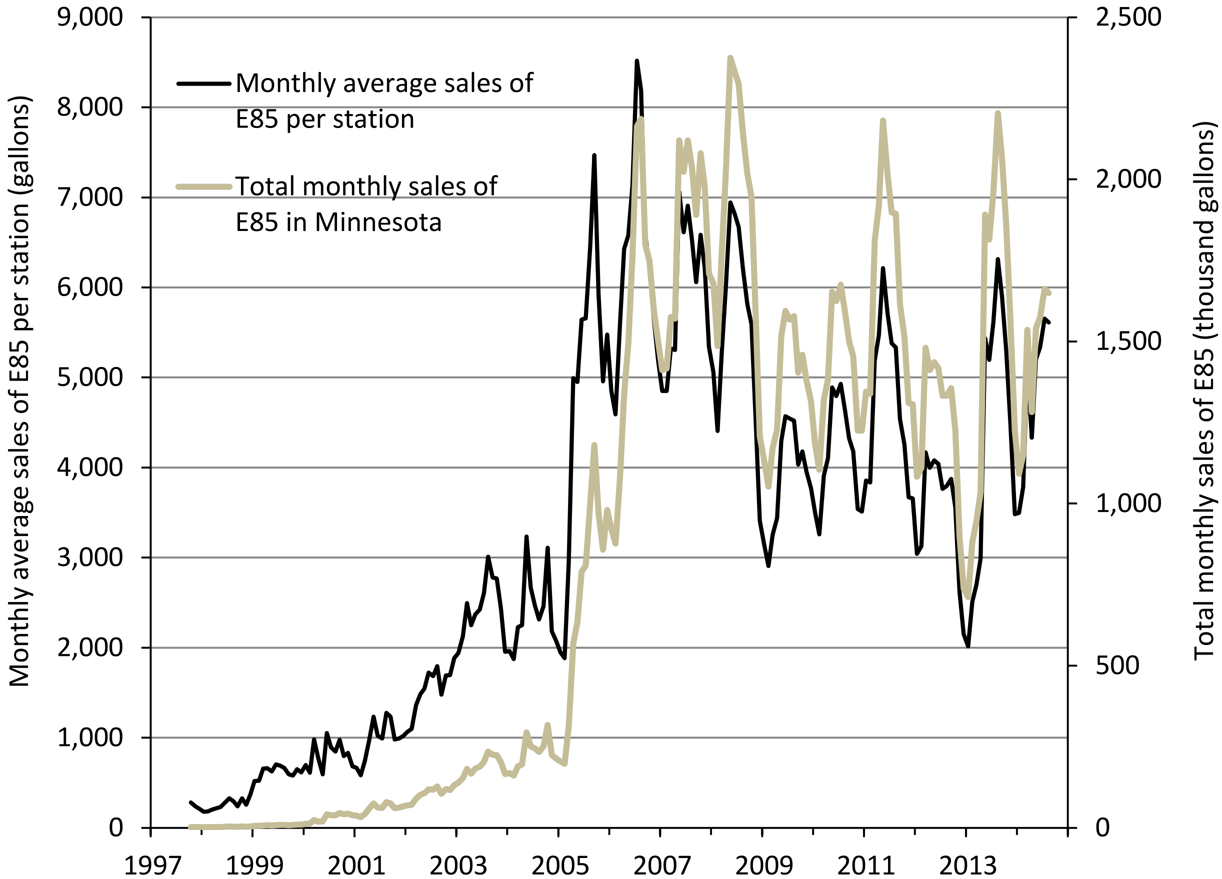

Figure 1: Average E85 Monthly Sales per Station

Notes: The data are from MN DoC (2014).

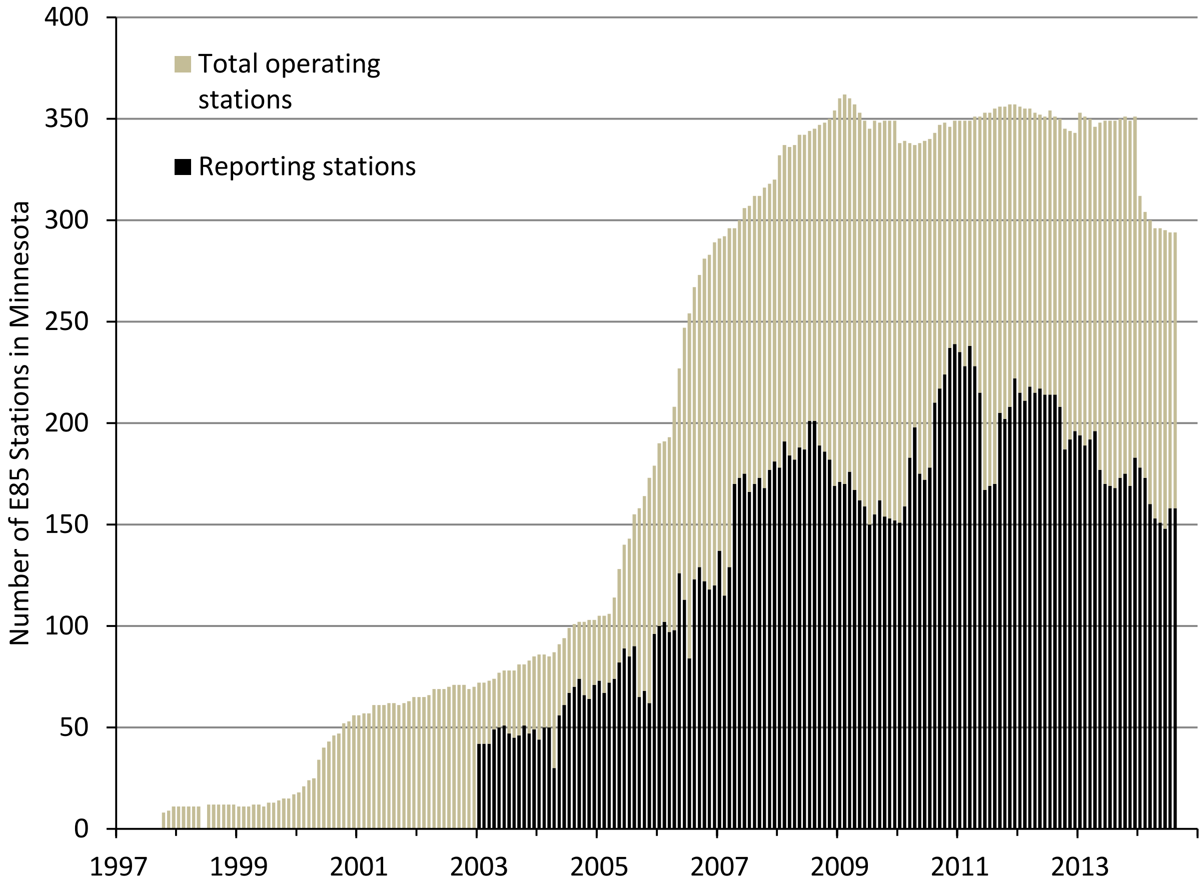

Figure 2: Number of Retail E85 Stations in Minnesota

Note: The data are from MN DoC (2014). MN DoC provides the number of reporting stations starting in January, 2003.

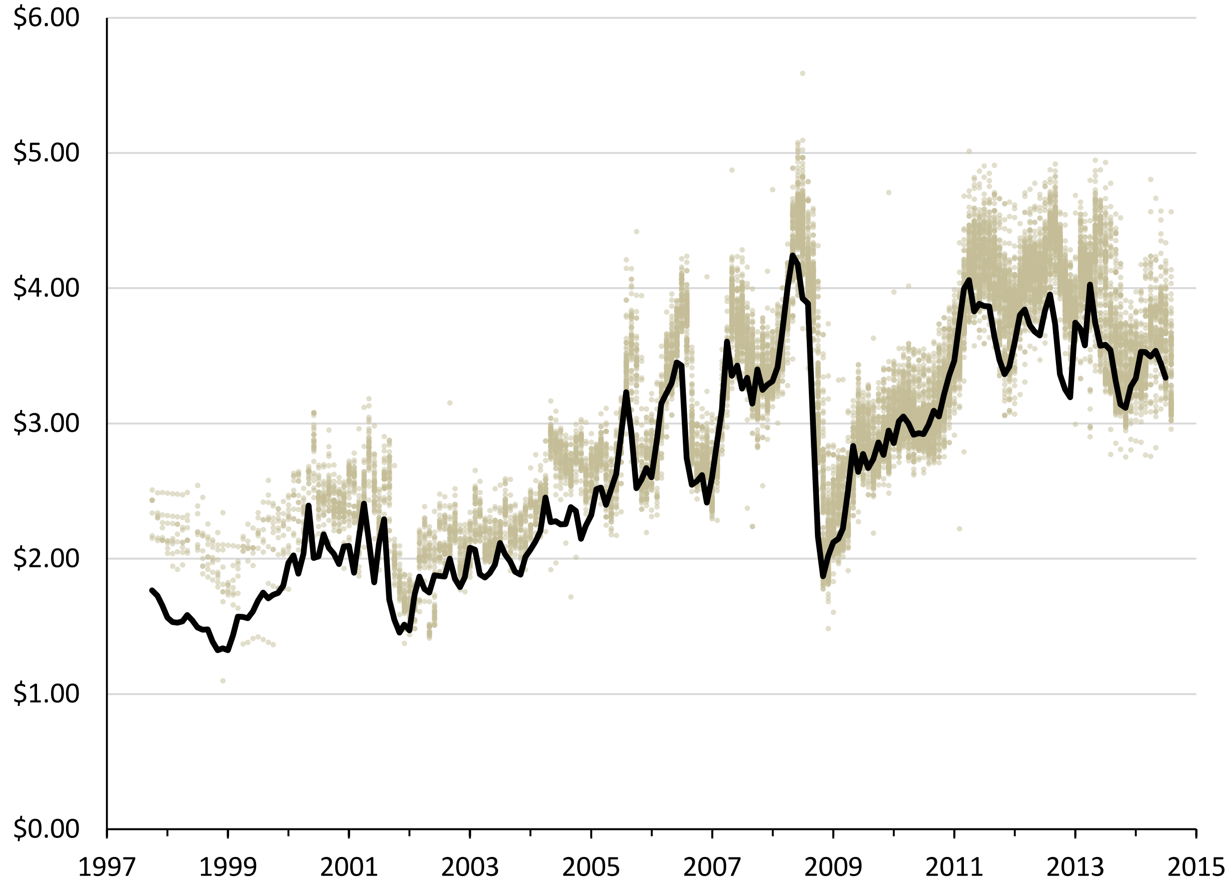

Figure 3: Retail E10 and Energy-Equivalent E85 Prices

Note: The data are from EIA (2014a, 2014b) and MN DoC (2014). Each dot represents an observation of the volume-weighted monthly average E85 price from an E85 station. The black line is the statewide average E10 price. E85 prices are measured in E10 energy equivalents, and all prices are in real August, 2014 dollars per gallon.

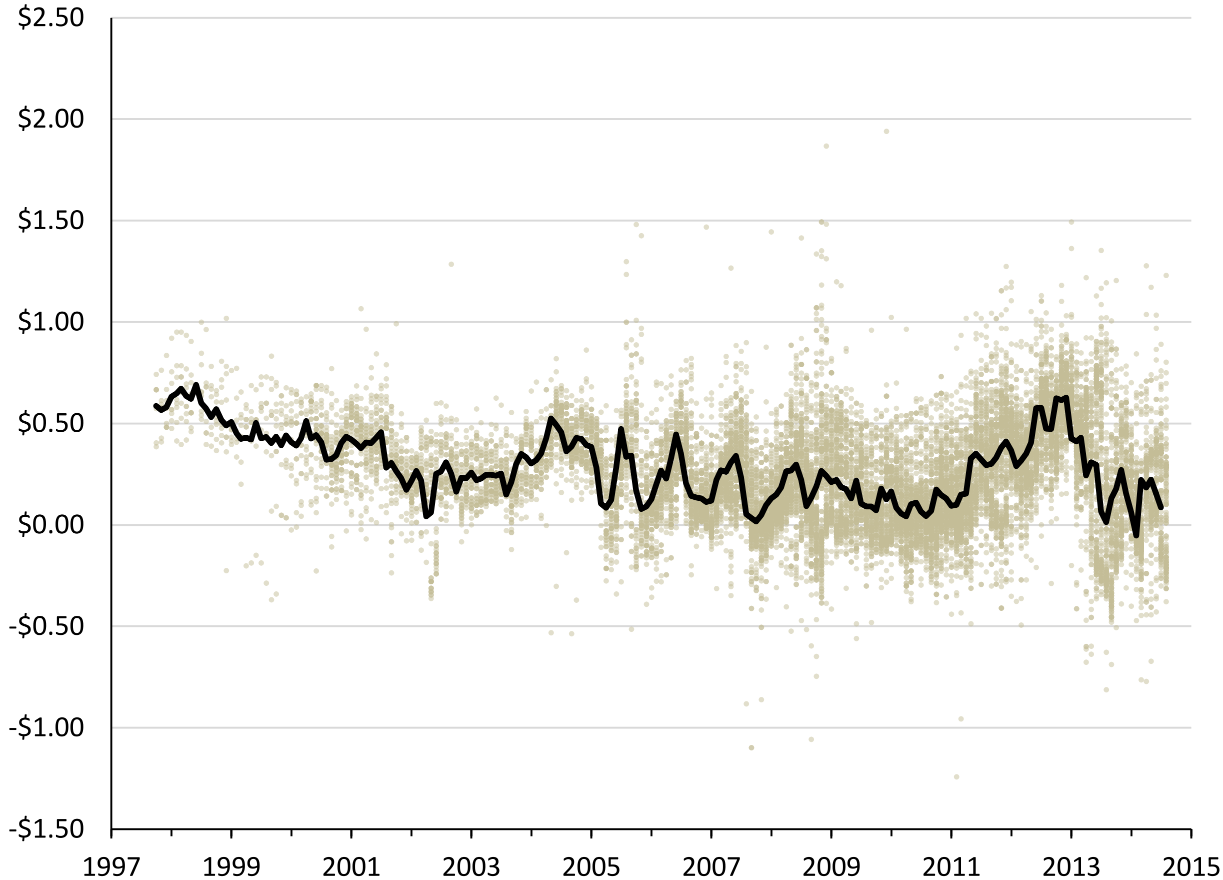

Figure 4: Energy-Equivalent Retail E85 Premiums

Note: The data are from EIA (2014a, 2014b) and MN DoC (2014). Each dot represents a monthly observation of the E85 premium from an E85 station. The black line is the average E85 premium from among reporting E85 stations. E85 prices are measured in E10 energy equivalents, and all prices are in real August, 2014 dollars per gallon.

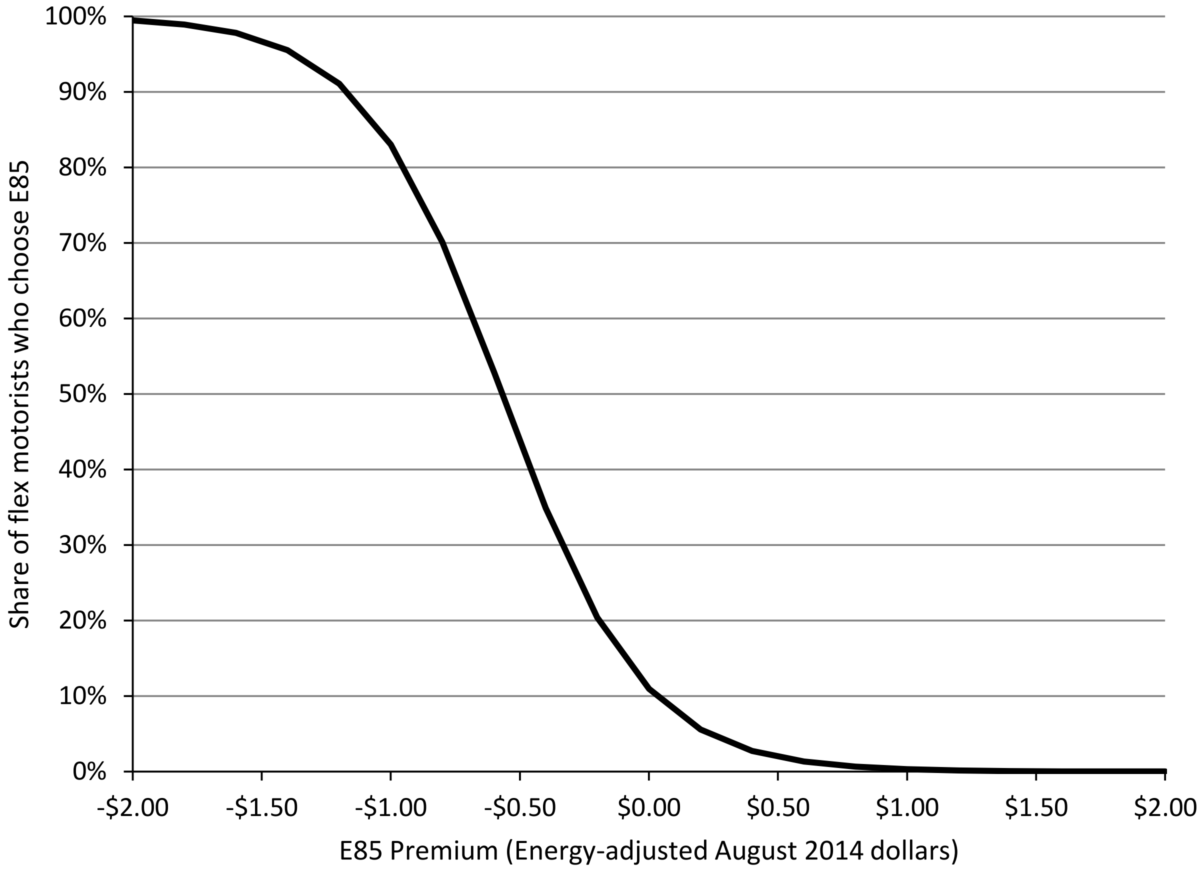

Figure 5: Cumulative Distribution Function (cdf) of Willingness to Pay for E85

Note: The cdf for the willingness to pay is calculated assuming a logistic distribution and using the estimates from Model 7, our preferred model, where $\mu = -0.569$ and $\sigma = 0.492$.

Table 1: Summary Statistics for Estimation Sample with All Stations

Variable	**Mean**	**Std. Dev**	**Min.**	**Max.**
Monthly retail E85 sales volume (gal)	3,771.46	3,400.51	3.17	38,955.77
Retail E85 price ($/gal).	3.535	0.613	1.484	5.591
Retail E10 price ($/gal)	3.314	0.510	1.869	4.241
Retail E85 premium ($/gal)	0.222	0.254	-1.242	1.941
Wholesale E85 price minus RIN ($/gal)	3.076	0.524	1.956	4.197
Wholesale E10 price ($/gal)	2.631	0.498	1.194	3.636
Wholesale corn price ($/bu)	5.601	1.419	2.913	8.295
Retail E85 station age (months)	60.448	42.723	2.000	203.000
Number of E85 stations in county	11.136	7.998	1.485	38.552
E85 stations per sq mi in county	0.054	0.072	0.004	0.276
All fuel stations per sq mi in county	0.183	0.281	0.005	1.143

Notes: Statistics are for 15,235 observations from 288 stations between 9/2006 and 8/2014. E85 prices and volumes are in E10 energy-equivalent terms. All prices are real August 2014 dollars.

Table 2: Summary Statistics for Estimation Sample with Identified Stations

Variable	**Mean**	**Std. Dev**	**Min.**	**Max.**
Monthly retail E85 sales volume (gal)	3,940.92	3,469.86	3.17	38,955.77
Retail E85 price ($/gal).	3.544	0.607	1.683	5.591
Retail E10 price ($/gal)	3.327	0.500	1.869	4.241
Retail E85 premium ($/gal)	0.217	0.251	-1.242	1.494
Wholesale E85 price minus RIN ($/gal)	3.071	0.518	1.956	4.197
Wholesale E10 price ($/gal)	2.641	0.487	1.194	3.636
Wholesale corn price ($/bu)	5.639	1.423	2.913	8.295
Retail E85 station age (months)	62.912	43.269	2.000	203.000
Number of E85 stations in county	11.249	8.044	1.485	38.552
E85 stations per sq mi in county	0.055	0.074	0.004	0.276
All fuel stations per sq mi in county	0.191	0.288	0.005	1.143

Notes: Statistics are for 13,941 observations from 246 stations between 9/2006 and 8/2014. E85 prices and volumes are in E10 energy-equivalent terms. All prices are real August 2014 dollars.

Table 3: Baseline Estimation Results

	Model 1	Model 2	Model 3	Model 4	Model 5
	All, OLS	Identified, OLS	All, 2SLS	Identified, 2SLS	Identified, 2SLS
Start	9/2006	9/2006	9/2006	9/2006	9/2006
Stations	288	246	288	246	246
Observations	15,235	13,941	15,235	13,941	13,941
Instruments	N/A	N/A	Simple	Simple	Complex
p_{it}	-0.899	-0.892	-1.281	-1.228	-1.206
	(0.020)	(0.020)	(0.033)	(0.032)	(0.032)
p_{it}^2	-0.084	-0.097	-0.136	-0.122	-0.165
	(0.028)	(0.031)	(0.050)	(0.052)	(0.052)
$\ln(\#stations)_{it}$	-0.045	-0.042	-0.057	-0.055	-0.054
	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)
μ	-1.452	-1.277	-1.207	-1.279	-1.002
	(0.404)	(0.381)	(0.321)	(0.383)	(0.285)
σ	1.671	1.636	1.215	1.272	1.226
	(0.080)	(0.086)	(0.050)	(0.060)	(0.058)

Notes: Estimates of the mean μ and the standard deviation σ of the willingness to pay are calculated by solving equations (8) and (9), and their standard errors are calculated using the delta method.

Table 4: Cubic Functions Allowing the Premium to Linearly Affect Market Size

	Model 2 Squared OLS	Model 5 Squared 2SLS	Model 6 Cubic OLS	Model 7 Cubic 2SLS
Start	9/2006	9/2006	9/2006	9/2006
Stations	246	246	246	246
Observations	13,941	13,941	13,941	13,941
p_{it}	-0.892	-1.206	-0.922	-1.167
	(0.020)	(0.032)	(0.020)	(0.032)
p_{it}^2	-0.097	-0.165	-0.520	-0.661
	(0.031)	(0.052)	(0.046)	(0.103)
p_{it}^3	NA	NA	0.565	0.636
	NA	NA	(0.045)	(0.114)
$\ln(\#stations)_{it}$	-0.042	-0.054	-0.041	-0.054
	(0.012)	(0.013)	(0.012)	(0.013)
μ	-1.277	-1.002	-0.649	-0.569
	(0.381)	(0.285)	(0.027)	(0.052)
σ	1.636	1.226	0.471	0.492
	(0.086)	(0.058)	(0.019)	(0.034)
BIC	7,737	8,573	7,587	8,551

Notes: For the squared models, estimates of μ and σ are calculated by solving equations (8) and (9), and their standards errors are calculated using the delta method. For the cubic models, estimates of μ and σ are calculated by solving equations (9) and (10) numerically, and their standards errors are calculated using Monte Carlo simulation.

Table 5: Relaxing Short-run Inelastic Demand

	Model 7	Model 8	Model 9	Model 10	Model 11
	$\eta = 0.00$	$\eta = -0.10$	$\eta = -0.30$	$\eta = -0.50$	η free
Start	9/2006	9/2006	9/2006	9/2006	9/2006
Stations	246	246	246	246	246
Observations	13,941	13,941	13,941	13,941	13,941
p_{it}	-1.167	-1.128	-1.050	-0.973	-1.494
	(0.032)	(0.033)	(0.033)	(0.034)	(0.035)
p_{it}^2	-0.661	-0.674	-0.699	-0.723	-0.557
	(0.103)	(0.103)	(0.105)	(0.106)	(0.103)
p_{it}^3	0.636	0.627	0.608	0.590	0.714
	(0.114)	(0.115)	(0.116)	(0.118)	(0.115)
$\ln(\#stations)_{it}$	-0.054	-0.056	-0.059	-0.061	-0.042
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
$\ln(P_e)_{it}$	0.000	-0.100	-0.300	-0.500	0.846
	NA	NA	NA	NA	(0.030)
μ	-0.569	-0.560	-0.541	-0.521	-0.621
	(0.052)	(0.052)	(0.050)	(0.049)	(0.047)
σ	0.492	0.500	0.515	0.529	0.413
	(0.034)	(0.035)	(0.038)	(0.041)	(0.028)

Notes: Estimates of μ and σ are calculated by solving equations (9) and (10) numerically, and their standard errors are calculated using Monte Carlo simulation.

Table 6: Different Estimation Samples (Start Dates)

	Model 12	Model 13	Model 14	Model 15
	Cubic 2SLS	Cubic 2SLS	Cubic 2SLS	Cubic 2SLS
Start	9/2004	9/2006	9/2008	9/2010
Stations	222	222	222	222
Observations	14,643	13,489	11,103	8,082
p_{it}	-1.115	-1.137	-1.153	-1.086
	(0.033)	(0.032)	(0.033)	(0.035)
p_{it}^2	-0.776	-0.718	-0.724	-0.619
	(0.105)	(0.100)	(0.099)	(0.102)
p_{it}^3	0.876	0.734	0.730	0.777
	(0.117)	(0.112)	(0.111)	(0.112)
$\ln(\#stations)_{it}$	-0.024	-0.052	-0.042	-0.030
	(0.011)	(0.012)	(0.015)	(0.018)
μ	-0.530	-0.548	-0.544	-0.593
	(0.040)	(0.045)	(0.044)	(0.043)
σ	0.431	0.466	0.469	0.415
	(0.020)	(0.026)	(0.026)	(0.024)

Notes: Estimates of μ and σ are calculated by solving equations (9) and (10) numerically, and their standard errors are calculated using Monte Carlo simulation.

Appendix A: Full estimation results from OLS and second stage regressions

Table A1: Model 1 Complete Results

Dependent variable	Log E85 sales volume
Number of observations	15,235
Number of stations	288
Method:	OLS
R-squared:	0.586

Variable	Estimate	Std. Error	*t*-statistic	p-value
Premium	-0.8993	0.0199	-45.1402	0.0000
Premium squared	-0.0837	0.0279	-3.0042	0.0027
Log E85 stations in county	-0.0447	0.0124	-3.5993	0.0003
Second month selling E85	-0.0878	0.0286	-3.0716	0.0021
Third month selling E85	-0.0682	0.0300	-2.2726	0.0231
Fourth month selling E85	-0.0506	0.0298	-1.6987	0.0894
Month 2	-0.0180	0.0125	-1.4382	0.1504
Month 3	0.1869	0.0128	14.5715	0.0000
Month 4	0.2720	0.0132	20.5501	0.0000
Month 5	0.4532	0.0139	32.5415	0.0000
Month 6	0.4316	0.0148	29.0973	0.0000
Month 7	0.4500	0.0158	28.4539	0.0000
Month 8	0.3932	0.0168	23.4465	0.0000
Month 9	0.2739	0.0173	15.8577	0.0000
Month 10	0.2272	0.0183	12.4387	0.0000
Month 11	0.0860	0.0195	4.4035	0.0000
Month 12	-0.0167	0.0207	-0.8053	0.4206
Year 2008	-0.0045	0.0247	-0.1820	0.8556
Year 2009	-0.4936	0.0430	-11.4902	0.0000
Year 2010	-0.4087	0.0616	-6.6312	0.0000
Year 2011	-0.1931	0.0805	-2.3991	0.0164
Year 2012	-0.2970	0.0995	-2.9833	0.0029
Year 2013	-0.5197	0.1186	-4.3802	0.0000
Year 2014	-0.6166	0.1376	-4.4809	0.0000
Station 1 trend	-0.0305	0.0062	-4.9218	0.0000
Station 2 trend	-0.0274	0.0024	-11.3761	0.0000
Station 3 trend	0.0097	0.0021	4.5342	0.0000
Station 5 trend	0.0097	0.0038	2.5715	0.0101
⋮	⋮	⋮	⋮	⋮
Station 514 trend	0.0024	0.0233	0.1043	0.9169
Station 516 trend	0.0054	0.0299	0.1800	0.8572
Station 518 trend	0.0481	0.0299	1.6077	0.1079
Station 519 trend	0.0051	0.0299	0.1696	0.8654

Table A2: Model 2 Complete Results

Dependent variable	Log E85 sales volume
Number of observations	13,941
Number of stations	246
Method:	OLS
R-squared:	0.602

Variable	Estimate	Std. Error	*t*-statistic	p-value
Premium	-0.8921	0.0204	-43.8055	0.0000
Premium squared	-0.0966	0.0306	-3.1607	0.0016
Log E85 stations in county	-0.0418	0.0124	-3.3618	0.0008
Second month selling E85	-0.0812	0.0287	-2.8339	0.0046
Third month selling E85	-0.0977	0.0298	-3.2763	0.0011
Fourth month selling E85	-0.0474	0.0299	-1.5856	0.1128
Month 2	-0.0209	0.0124	-1.6816	0.0926
Month 3	0.1887	0.0128	14.7597	0.0000
Month 4	0.2601	0.0132	19.6504	0.0000
Month 5	0.4374	0.0140	31.2862	0.0000
Month 6	0.4173	0.0149	27.9615	0.0000
Month 7	0.4372	0.0160	27.3973	0.0000
Month 8	0.3788	0.0170	22.3069	0.0000
Month 9	0.2670	0.0176	15.2067	0.0000
Month 10	0.2188	0.0186	11.7532	0.0000
Month 11	0.0899	0.0200	4.4974	0.0000
Month 12	-0.0219	0.0212	-1.0347	0.3008
Year 2008	-0.0110	0.0257	-0.4268	0.6695
Year 2009	-0.4981	0.0445	-11.1832	0.0000
Year 2010	-0.4117	0.0638	-6.4535	0.0000
Year 2011	-0.2011	0.0833	-2.4140	0.0158
Year 2012	-0.3050	0.1030	-2.9623	0.0031
Year 2013	-0.5216	0.1227	-4.2522	0.0000
Year 2014	-0.6198	0.1422	-4.3572	0.0000
Station 2 trend	-0.0273	0.0024	-11.4740	0.0000
Station 3 trend	0.0096	0.0021	4.5217	0.0000
Station 5 trend	0.0097	0.0037	2.6453	0.0082
Station 6 trend	-0.0019	0.0020	-0.9381	0.3482
⋮	⋮	⋮	⋮	⋮
Station 501 trend	0.0113	0.0154	0.7326	0.4638
Station 514 trend	0.0014	0.0221	0.0626	0.9501
Station 516 trend	0.0063	0.0285	0.2215	0.8247
Station 518 trend	0.0491	0.0285	1.7258	0.0844

Table A3: Model 3 Second Stage Results

Dependent variable	Log E85 sales volume
Number of observations	15,235
Number of stations	288
Method:	2SLS (simple instruments)
Weak instruments *F*-statistic (p-value):	1266.1 (0.000)

Variable	Estimate	Std. Error	*t*-statistic	p-value
Premium	-1.2810	0.0332	-38.6211	0.0000
Premium squared	-0.1355	0.0503	-2.6926	0.0071
Log E85 stations in county	-0.0572	0.0128	-4.4815	0.0000
Second month selling E85	-0.0920	0.0294	-3.1328	0.0017
Third month selling E85	-0.0669	0.0309	-2.1681	0.0302
Fourth month selling E85	-0.0458	0.0307	-1.4934	0.1353
Month 2	-0.0368	0.0129	-2.8597	0.0042
Month 3	0.1580	0.0133	11.9152	0.0000
Month 4	0.2600	0.0136	19.0540	0.0000
Month 5	0.4380	0.0144	30.5057	0.0000
Month 6	0.4373	0.0153	28.6109	0.0000
Month 7	0.4727	0.0163	29.0087	0.0000
Month 8	0.3830	0.0173	22.2044	0.0000
Month 9	0.2338	0.0178	13.1074	0.0000
Month 10	0.1967	0.0188	10.4483	0.0000
Month 11	0.0759	0.0201	3.7779	0.0002
Month 12	-0.0098	0.0213	-0.4627	0.6436
Year 2008	0.0053	0.0254	0.2077	0.8355
Year 2009	-0.4872	0.0442	-11.0305	0.0000
Year 2010	-0.4296	0.0634	-6.7802	0.0000
Year 2011	-0.1615	0.0828	-1.9507	0.0511
Year 2012	-0.1746	0.1025	-1.7036	0.0884
Year 2013	-0.4675	0.1220	-3.8316	0.0001
Year 2014	-0.6242	0.1415	-4.4118	0.0000
Station 1 trend	-0.0283	0.0064	-4.4444	0.0000
Station 2 trend	-0.0218	0.0025	-8.7268	0.0000
Station 3 trend	0.0106	0.0022	4.8520	0.0000
Station 5 trend	0.0093	0.0039	2.4047	0.0162
⋮	⋮	⋮	⋮	⋮
Station 514 trend	-0.0129	0.0239	-0.5408	0.5887
Station 516 trend	0.0064	0.0307	0.2075	0.8356
Station 518 trend	0.0546	0.0308	1.7751	0.0759
Station 519 trend	0.0075	0.0307	0.2433	0.8078

Table A4: Model 4 Second Stage Results

Dependent variable	Log E85 sales volume
Number of observations	13,941
Number of stations	246
Method:	2SLS (simple instruments)
Weak instruments *F*-statistic (p-value):	1253.9 (0.000)

Variable	Estimate	Std. Error	*t*-statistic	p-value
Premium	-1.2277	0.0324	-37.8840	0.0000
Premium squared	-0.1216	0.0518	-2.3500	0.0188
Log E85 stations in county	-0.0547	0.0127	-4.2935	0.0000
Second month selling E85	-0.0844	0.0293	-2.8832	0.0039
Third month selling E85	-0.1014	0.0305	-3.3272	0.0009
Fourth month selling E85	-0.0423	0.0306	-1.3838	0.1664
Month 2	-0.0373	0.0127	-2.9266	0.0034
Month 3	0.1636	0.0131	12.4530	0.0000
Month 4	0.2500	0.0136	18.4380	0.0000
Month 5	0.4229	0.0143	29.5220	0.0000
Month 6	0.4204	0.0153	27.5139	0.0000
Month 7	0.4549	0.0163	27.8500	0.0000
Month 8	0.3694	0.0174	21.2776	0.0000
Month 9	0.2329	0.0180	12.9282	0.0000
Month 10	0.1936	0.0191	10.1559	0.0000
Month 11	0.0814	0.0204	3.9860	0.0001
Month 12	-0.0163	0.0217	-0.7509	0.4527
Year 2008	-0.0026	0.0263	-0.1007	0.9198
Year 2009	-0.4908	0.0455	-10.7800	0.0000
Year 2010	-0.4254	0.0652	-6.5255	0.0000
Year 2011	-0.1689	0.0852	-1.9834	0.0473
Year 2012	-0.1940	0.1054	-1.8411	0.0656
Year 2013	-0.4703	0.1254	-3.7507	0.0002
Year 2014	-0.6174	0.1454	-4.2469	0.0000
Station 2 trend	-0.0228	0.0025	-9.2665	0.0000
Station 3 trend	0.0104	0.0022	4.7616	0.0000
Station 5 trend	0.0092	0.0037	2.4793	0.0132
Station 6 trend	-0.0012	0.0021	-0.5861	0.5578
⋮	⋮	⋮	⋮	⋮
Station 501 trend	0.0120	0.0157	0.7597	0.4474
Station 514 trend	-0.0119	0.0226	-0.5266	0.5984
Station 516 trend	0.0070	0.0291	0.2422	0.8086
Station 518 trend	0.0544	0.0291	1.8712	0.0613

Table A5: Model 5 Second Stage Results

Dependent variable	Log E85 sales volume
Number of observations	13,941
Number of stations	246
Method:	2SLS (complex instruments)
Weak instruments *F*-statistic (p-value):	244.96 (0.000)

Variable	Estimate	Std. Error	*t*-statistic	p-value
Premium	-1.2060	0.0320	-37.6328	0.0000
Premium squared	-0.1651	0.0519	-3.1810	0.0015
Log E85 stations in county	-0.0543	0.0127	-4.2625	0.0000
Second month selling E85	-0.0844	0.0293	-2.8804	0.0040
Third month selling E85	-0.1014	0.0305	-3.3251	0.0009
Fourth month selling E85	-0.0424	0.0306	-1.3885	0.1650
Month 2	-0.0376	0.0127	-2.9504	0.0032
Month 3	0.1631	0.0131	12.4169	0.0000
Month 4	0.2493	0.0136	18.3864	0.0000
Month 5	0.4223	0.0143	29.4859	0.0000
Month 6	0.4198	0.0153	27.4782	0.0000
Month 7	0.4550	0.0163	27.8580	0.0000
Month 8	0.3698	0.0174	21.2995	0.0000
Month 9	0.2330	0.0180	12.9367	0.0000
Month 10	0.1937	0.0191	10.1600	0.0000
Month 11	0.0811	0.0204	3.9695	0.0001
Month 12	-0.0164	0.0217	-0.7545	0.4505
Year 2008	-0.0032	0.0263	-0.1226	0.9024
Year 2009	-0.4919	0.0455	-10.8055	0.0000
Year 2010	-0.4265	0.0652	-6.5422	0.0000
Year 2011	-0.1713	0.0852	-2.0110	0.0443
Year 2012	-0.1952	0.1054	-1.8525	0.0640
Year 2013	-0.4718	0.1254	-3.7624	0.0002
Year 2014	-0.6197	0.1454	-4.2627	0.0000
Station 2 trend	-0.0226	0.0025	-9.1599	0.0000
Station 3 trend	0.0103	0.0022	4.7552	0.0000
Station 5 trend	0.0093	0.0037	2.4856	0.0129
Station 6 trend	-0.0012	0.0021	-0.5695	0.5690
⋮	⋮	⋮	⋮	⋮
Station 501 trend	0.0120	0.0157	0.7628	0.4456
Station 514 trend	-0.0119	0.0226	-0.5242	0.6001
Station 516 trend	0.0071	0.0291	0.2430	0.8080
Station 518 trend	0.0544	0.0291	1.8706	0.0614

Table A6: Model 6 Complete Results

Dependent variable	Log E85 sales volume
Number of observations	13,941
Number of stations	246
Method:	OLS
R-squared:	0.607

Variable	Estimate	Std. Error	*t*-statistic	p-value
Premium	-0.9218	0.0204	-45.2070	0.0000
Premium squared	-0.5197	0.0456	-11.3905	0.0000
Premium cubed	0.5652	0.0455	12.4301	0.0000
Log E85 stations in county	-0.0409	0.0124	-3.3082	0.0009
Second month selling E85	-0.0842	0.0285	-2.9569	0.0031
Third month selling E85	-0.1026	0.0297	-3.4601	0.0005
Fourth month selling E85	-0.0485	0.0297	-1.6330	0.1025
Month 2	-0.0199	0.0124	-1.6112	0.1071
Month 3	0.1886	0.0127	14.8361	0.0000
Month 4	0.2615	0.0132	19.8691	0.0000
Month 5	0.4392	0.0139	31.5983	0.0000
Month 6	0.4185	0.0148	28.1996	0.0000
Month 7	0.4398	0.0159	27.7151	0.0000
Month 8	0.3791	0.0169	22.4529	0.0000
Month 9	0.2665	0.0175	15.2633	0.0000
Month 10	0.2139	0.0185	11.5542	0.0000
Month 11	0.0871	0.0199	4.3807	0.0000
Month 12	-0.0246	0.0211	-1.1643	0.2443
Year 2008	-0.0145	0.0256	-0.5671	0.5707
Year 2009	-0.5065	0.0443	-11.4366	0.0000
Year 2010	-0.4285	0.0634	-6.7545	0.0000
Year 2011	-0.2114	0.0828	-2.5519	0.0107
Year 2012	-0.3047	0.1024	-2.9761	0.0029
Year 2013	-0.5321	0.1220	-4.3625	0.0000
Year 2014	-0.6467	0.1414	-4.5722	0.0000
Station 2 trend	-0.0292	0.0024	-12.3193	0.0000
Station 3 trend	0.0100	0.0021	4.7294	0.0000
Station 5 trend	0.0098	0.0036	2.6951	0.0070
Station 6 trend	-0.0015	0.0020	-0.7546	0.4505
⋮	⋮	⋮	⋮	⋮
Station 501 trend	0.0113	0.0153	0.7381	0.4605
Station 514 trend	-0.0019	0.0220	-0.0844	0.9328
Station 516 trend	0.0065	0.0283	0.2282	0.8195
Station 518 trend	0.0508	0.0283	1.7938	0.0729

Table A7: Model 7 Second Stage Results

Dependent variable	Log E85 sales volume
Number of observations	13,941
Number of stations	246
Method:	2SLS (complex instruments)
Weak instruments *F*-statistic (p-value):	245.0 (0.000)

Variable	Estimate	Std. Error	*t*-statistic	p-value
Premium	-1.1665	0.0325	-35.9477	0.0000
Premium squared	-0.6614	0.1030	-6.4241	0.0000
Premium cubed	0.6358	0.1144	5.5590	0.0000
Log E85 stations in county	-0.0542	0.0126	-4.3087	0.0000
Second month selling E85	-0.0860	0.0289	-2.9709	0.0030
Third month selling E85	-0.1024	0.0301	-3.4002	0.0007
Fourth month selling E85	-0.0429	0.0302	-1.4205	0.1555
Month 2	-0.0366	0.0126	-2.9059	0.0037
Month 3	0.1636	0.0130	12.6033	0.0000
Month 4	0.2505	0.0134	18.6983	0.0000
Month 5	0.4237	0.0142	29.9407	0.0000
Month 6	0.4228	0.0151	27.9908	0.0000
Month 7	0.4567	0.0161	28.2992	0.0000
Month 8	0.3728	0.0172	21.7222	0.0000
Month 9	0.2340	0.0178	13.1477	0.0000
Month 10	0.1945	0.0188	10.3268	0.0000
Month 11	0.0802	0.0202	3.9752	0.0001
Month 12	-0.0180	0.0214	-0.8424	0.3996
Year 2008	-0.0052	0.0260	-0.2017	0.8401
Year 2009	-0.4953	0.0450	-11.0099	0.0000
Year 2010	-0.4326	0.0644	-6.7144	0.0000
Year 2011	-0.1739	0.0841	-2.0671	0.0387
Year 2012	-0.1931	0.1041	-1.8552	0.0636
Year 2013	-0.4701	0.1239	-3.7941	0.0001
Year 2014	-0.6263	0.1436	-4.3603	0.0000
Station 2 trend	-0.0247	0.0025	-10.0139	0.0000
Station 3 trend	0.0104	0.0021	4.8560	0.0000
Station 5 trend	0.0092	0.0037	2.4902	0.0128
Station 6 trend	-0.0010	0.0020	-0.5023	0.6155
⋮	⋮	⋮	⋮	⋮
Station 501 trend	0.0118	0.0155	0.7598	0.4473
Station 514 trend	-0.0137	0.0224	-0.6145	0.5389
Station 516 trend	0.0071	0.0287	0.2477	0.8043
Station 518 trend	0.0554	0.0287	1.9269	0.0540

Table A8: Model 11 Second Stage Results

Dependent variable	Log E85 sales volume
Number of observations	13,941
Number of stations	246
Method:	2SLS (complex instruments)

Variable	Estimate	Std. Error	*t*-statistic	p-value
Premium	-1.4942	0.0345	-43.2824	0.0000
Premium squared	-0.5565	0.1032	-5.3914	0.0000
Premium cubed	0.7141	0.1146	6.2298	0.0000
Log E85 price	0.8457	0.0300	28.2204	0.0000
Log E85 stations in county	-0.0420	0.0126	-3.3321	0.0009
Second month selling E85	-0.0791	0.0290	-2.7269	0.0064
Third month selling E85	-0.0884	0.0302	-2.9293	0.0034
Fourth month selling E85	-0.0330	0.0302	-1.0907	0.2754
Month 2	-0.0624	0.0126	-4.9344	0.0000
Month 3	0.1120	0.0131	8.5267	0.0000
Month 4	0.1906	0.0136	14.0281	0.0000
Month 5	0.3370	0.0145	23.2328	0.0000
Month 6	0.3560	0.0153	23.2396	0.0000
Month 7	0.4104	0.0163	25.2477	0.0000
Month 8	0.3384	0.0172	19.6301	0.0000
Month 9	0.2017	0.0179	11.2868	0.0000
Month 10	0.2238	0.0189	11.8427	0.0000
Month 11	0.1669	0.0205	8.1579	0.0000
Month 12	0.0942	0.0218	4.3136	0.0000
Year 2008	0.0818	0.0262	3.1221	0.0018
Year 2009	-0.0603	0.0476	-1.2660	0.2055
Year 2010	-0.0180	0.0662	-0.2715	0.7860
Year 2011	0.2047	0.0854	2.3976	0.0165
Year 2012	0.3244	0.1059	3.0636	0.0022
Year 2013	0.1948	0.1263	1.5422	0.1230
Year 2014	0.1870	0.1468	1.2743	0.2026
Station 2 trend	-0.0357	0.0025	-14.2976	0.0000
Station 3 trend	0.0001	0.0022	0.0596	0.9525
Station 5 trend	-0.0006	0.0037	-0.1682	0.8664
Station 6 trend	-0.0119	0.0021	-5.7193	0.0000
⋮	⋮	⋮	⋮	⋮
Station 501 trend	0.0038	0.0156	0.2424	0.8084
Station 514 trend	-0.0207	0.0224	-0.9253	0.3548
Station 516 trend	0.0001	0.0288	0.0037	0.9971
Station 518 trend	0.0488	0.0288	1.6951	0.0901

Appendix B: First stage estimation results

Table B1: Model 3 First Stage Results (Simple Instruments and All Observations)

Dependent variable	E85 premium
Number of observations	15,235
Number of stations	288
R-squared:	0.7057
F-statistic (p-value):	57.99 (0.000)

Variable	Estimate	Std. Error	*t*-statistic	p-value
Log E85 stations in county	-0.0097	0.0056	-1.7303	0.0836
Second month selling E85	0.0037	0.0131	0.2802	0.7794
Third month selling E85	0.0101	0.0135	0.7476	0.4547
Fourth month selling E85	0.0131	0.0134	0.9783	0.3279
Wholesale E85 * All density	0.0240	0.0071	3.3770	0.0007
Wholesale E10 * All density	-0.0118	0.0079	-1.4963	0.1346
Wholesale E85 * E85 density	-0.0555	0.0059	-9.3605	0.0000
Wholesale E10 * E85 density	0.0610	0.0067	9.1083	0.0000
Corn price	0.0271	0.0020	13.5457	0.0000
Lag log E85 price	0.0504	0.0216	2.3337	0.0196
Lag log E85 quantity	0.0049	0.0039	1.2582	0.2083
Lag E85 premium	0.5486	0.0099	55.1389	0.0000
Month 2	-0.0185	0.0057	-3.2485	0.0012
Month 3	-0.0169	0.0059	-2.8729	0.0041
Month 4	0.0316	0.0062	5.1411	0.0000
Month 5	0.0131	0.0066	1.9883	0.0468
Month 6	0.0547	0.0072	7.6166	0.0000
Month 7	0.0527	0.0075	6.9905	0.0000
Month 8	-0.0610	0.0080	-7.6725	0.0000
Month 9	-0.1027	0.0081	-12.6339	0.0000
Month 10	-0.0558	0.0086	-6.4763	0.0000
Month 11	-0.0432	0.0092	-4.7042	0.0000
Month 12	-0.0380	0.0096	-3.9746	0.0001
Year 2008	-0.0701	0.0115	-6.1018	0.0000
Year 2009	-0.1172	0.0206	-5.6884	0.0000
Year 2010	-0.2108	0.0285	-7.4027	0.0000
Year 2011	-0.2824	0.0370	-7.6241	0.0000
Year 2012	-0.2180	0.0458	-4.7650	0.0000
Year 2013	-0.3154	0.0544	-5.7986	0.0000
Year 2014	-0.4310	0.0631	-6.8253	0.0000
Station 1 trend	0.0093	0.0028	3.3082	0.0009
Station 2 trend	0.0121	0.0011	10.9988	0.0000
⋮	⋮	⋮	⋮	⋮
Station 519 trend	0.0065	0.0135	0.4816	0.6301

Table B2: Model 4 First Stage Results (Simple Instruments and Identified Observations)

Dependent variable	E85 premium
Number of observations	13,941
Number of stations	246
R-squared:	0.8379
F-statistic (p-value):	132.9 (0.000)

Variable	Estimate	Std. Error	*t*-statistic	p-value
Log E85 stations in county	-0.0125	0.0057	-2.1953	0.0282
Second month selling E85	0.0132	0.0134	0.9831	0.3256
Third month selling E85	0.0005	0.0137	0.0333	0.9734
Fourth month selling E85	0.0234	0.0137	1.7040	0.0884
Wholesale E85 * All density	0.0306	0.0070	4.3632	0.0000
Wholesale E10 * All density	-0.0207	0.0079	-2.6331	0.0085
Wholesale E85 * E85 density	-0.0606	0.0059	-10.3175	0.0000
Wholesale E10 * E85 density	0.0684	0.0067	10.2020	0.0000
Corn price	0.0276	0.0020	13.7413	0.0000
Lag log E85 price	0.0609	0.0221	2.7592	0.0058
Lag log E85 quantity	0.0135	0.0041	3.2801	0.0010
Lag E85 premium	0.5643	0.0103	54.8256	0.0000
Month 2	-0.0193	0.0058	-3.3483	0.0008
Month 3	-0.0183	0.0060	-3.0610	0.0022
Month 4	0.0295	0.0063	4.7049	0.0000
Month 5	0.0040	0.0067	0.5951	0.5518
Month 6	0.0477	0.0074	6.4914	0.0000
Month 7	0.0461	0.0077	5.9625	0.0000
Month 8	-0.0669	0.0082	-8.1793	0.0000
Month 9	-0.1072	0.0084	-12.7362	0.0000
Month 10	-0.0599	0.0089	-6.6994	0.0000
Month 11	-0.0479	0.0096	-4.9930	0.0000
Month 12	-0.0451	0.0100	-4.5118	0.0000
Year 2008	-0.0734	0.0122	-6.0289	0.0000
Year 2009	-0.1092	0.0217	-5.0247	0.0000
Year 2010	-0.2019	0.0300	-6.7213	0.0000
Year 2011	-0.2796	0.0390	-7.1614	0.0000
Year 2012	-0.2159	0.0482	-4.4750	0.0000
Year 2013	-0.3078	0.0573	-5.3682	0.0000
Year 2014	-0.4209	0.0666	-6.3242	0.0000
Station 2 trend	0.0121	0.0011	10.8833	0.0000
Station 3 trend	0.0069	0.0010	6.9962	0.0000
⋮	⋮	⋮	⋮	⋮
Station 518 trend	0.0063	0.0131	0.4857	0.6272

Table B3: Model 5 First Stage Results (Complex Instruments and Identified Observations)

Dependent variable	E85 premium
Number of observations	13,941
Number of stations	246
R-squared:	0.8399
F-statistic (p-value):	126.3 (0.000)

Variable	Estimate	Std. Error	*t*-statistic	p-value
Log E85 stations in county	-0.0107	0.0057	-1.8760	0.0607
Second month selling E85	0.0110	0.0134	0.8191	0.4127
Third month selling E85	0.0005	0.0137	0.0357	0.9715
Fourth month selling E85	0.0228	0.0137	1.6653	0.0959
Wholesale E85 * All density	0.0064	0.0124	0.5188	0.6039
Wholesale E10 * All density	-0.0071	0.0137	-0.5133	0.6078
Wholesale E85 * E85 density	-0.0115	0.0127	-0.9040	0.3660
Wholesale E10 * E85 density	0.0129	0.0142	0.9055	0.3652
Wholesale E85 * BP	0.0743	0.0416	1.7844	0.0744
Wholesale E85 * Clark	0.0738	0.0646	1.1413	0.2538
⋮	⋮	⋮	⋮	⋮
Wholesale E85 * Tesoro	0.1294	0.0581	2.2272	0.0260
Wholesale E10 * BP	-0.1495	0.0466	-3.2075	0.0013
Wholesale E10 * Clark	-0.0462	0.0717	-0.6442	0.5195
⋮	⋮	⋮	⋮	⋮
Wholesale E10 * Tesoro	-0.1388	0.0604	-2.2971	0.0216
Wholesale E85 * Log distance to blending terminal	0.0186	0.0088	2.1099	0.0349
Wholesale E10 * Log distance to blending terminal	-0.0294	0.0098	-3.0025	0.0027
Corn price	0.0281	0.0021	13.3071	0.0000
Lag E85 premium	0.5529	0.0106	52.1257	0.0000
Lag log E85 price	0.0773	0.0238	3.2491	0.0012
Lag log E85 quantity	0.0132	0.0041	3.2274	0.0013
Month 2	-0.0180	0.0058	-3.1246	0.0018
Month 3	-0.0174	0.0060	-2.9193	0.0035
Month 4	0.0303	0.0063	4.8446	0.0000
Month 5	0.0059	0.0067	0.8689	0.3849
Month 6	0.0478	0.0073	6.5202	0.0000
Month 7	0.0460	0.0077	5.9551	0.0000
Month 8	-0.0670	0.0082	-8.2068	0.0000
Month 9	-0.1084	0.0084	-12.9158	0.0000
Month 10	-0.0635	0.0089	-7.0929	0.0000
Month 11	-0.0523	0.0096	-5.4507	0.0000
Month 12	-0.0478	0.0100	-4.7834	0.0000
Year 2008	-0.0767	0.0122	-6.2887	0.0000

Year 2009	-0.1107	0.0217	-5.1005	0.0000
Year 2010	-0.2070	0.0300	-6.9001	0.0000
Year 2011	-0.2886	0.0390	-7.3985	0.0000
Year 2012	-0.2229	0.0482	-4.6237	0.0000
Year 2013	-0.3149	0.0573	-5.4978	0.0000
Year 2014	-0.4298	0.0664	-6.4703	0.0000
Station 2 trend	0.0126	0.0011	11.3763	0.0000
Station 3 trend	0.0074	0.0010	7.2600	0.0000
⋮	⋮	⋮	⋮	⋮
Station 518 trend	0.0088	0.0130	0.6738	0.5005

Table B4: Model 8 First Stage Results (Complex Instruments for Log E85 price)

Dependent variable	Log E85 price
Number of observations	13,941
Number of stations	246
R-squared:	0.989
F-statistic (p-value):	2.373e+04 (0.000)

Variable	Estimate	Std. Error	*t*-statistic	p-value
Log E85 stations in county	-0.0048	0.0017	-2.7608	0.0058
Second month selling E85	0.0005	0.0041	0.1132	0.9099
Third month selling E85	0.0018	0.0041	0.4448	0.6565
Fourth month selling E85	0.0106	0.0041	2.5539	0.0107
Wholesale E85 * All density	-0.0012	0.0037	-0.3333	0.7389
Wholesale E10 * All density	0.0019	0.0042	0.4448	0.6565
Wholesale E85 * E85 density	-0.0003	0.0038	-0.0705	0.9438
Wholesale E10 * E85 density	0.0022	0.0043	0.5125	0.6083
Wholesale E85 * BP	0.0333	0.0126	2.6418	0.0083
Wholesale E85 * Clark	0.0445	0.0196	2.2719	0.0231
⋮	⋮	⋮	⋮	⋮
Wholesale E85 * Tesoro	0.0576	0.0176	3.2717	0.0011
Wholesale E10 * BP	0.2257	0.0141	15.9796	0.0000
Wholesale E10 * Clark	0.2238	0.0217	10.3083	0.0000
⋮	⋮	⋮	⋮	⋮
Wholesale E10 * Tesoro	0.2265	0.0183	12.3767	0.0000
Wholesale E85 * Log distance to blending terminal	0.0047	0.0027	1.7386	0.0821
Wholesale E10 * Log distance to blending terminal	-0.0083	0.0030	-2.7910	0.0053
Corn price	0.0030	0.0006	4.6825	0.0000
Lag E85 premium	0.0928	0.0032	28.8632	0.0000
Lag log E85 price	0.2286	0.0072	31.6933	0.0000
Lag log E85 quantity	-0.0014	0.0012	-1.1068	0.2684
Month 2	-0.0128	0.0017	-7.3439	0.0000
Month 3	-0.0185	0.0018	-10.2086	0.0000
Month 4	-0.0164	0.0019	-8.6716	0.0000
Month 5	-0.0178	0.0020	-8.7440	0.0000
Month 6	-0.0078	0.0022	-3.5034	0.0005
Month 7	-0.0198	0.0023	-8.4517	0.0000
Month 8	-0.0531	0.0025	-21.4643	0.0000
Month 9	-0.0476	0.0025	-18.7170	0.0000
Month 10	-0.0344	0.0027	-12.6808	0.0000
Month 11	-0.0318	0.0029	-10.9368	0.0000
Month 12	-0.0257	0.0030	-8.5048	0.0000
Year 2008	-0.0616	0.0037	-16.6643	0.0000

Year 2009	-0.0345	0.0066	-5.2518	0.0000
Year 2010	-0.0614	0.0091	-6.7601	0.0000
Year 2011	-0.1131	0.0118	-9.5720	0.0000
Year 2012	-0.0971	0.0146	-6.6471	0.0000
Year 2013	-0.1197	0.0174	-6.8998	0.0000
Year 2014	-0.1631	0.0201	-8.1030	0.0000
Station 2 trend	0.0039	0.0003	11.5512	0.0000
Station 3 trend	0.0030	0.0003	9.7260	0.0000
⋮	⋮	⋮	⋮	⋮
Station 518 trend	0.0037	0.0039	0.9407	0.3469