NBER WORKING PAPER SERIES

THE DEMAND FOR ETHANOL AS A GASOLINE SUBSTITUTE

Soren T. Anderson

Working Paper 16371
http://www.nber.org/papers/w16371

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2010

For helpful comments and suggestions, I thank Matias Busso, Brian Cadena, Lucas Davis, Alex Farrell, Meredith Fowlie, Ben Keys, Brian Kovak, Erin Mansur, Michael Moore, Alex Resch, Stephen Salant, Jim Sallee, Gary Solon, Roger von Haefen, Sarah West, and seminar participants at the University of Michigan, University of California Energy Institute, NBER Summer Institute, Michigan State University, Triangle Resource and Environmental Economics Seminar, MIT, and NBER. I thank the Minnesota Department of Commerce and the American Lung Association of Minnesota for providing retail ethanol price and sales volume data, the Minnesota Department of Public Safety for providing vehicle registration data, and the U.S. Department of Transportation for providing vehicle sales data. I thank the University of Michigan's Center for Local, State, and Urban Policy and Rackham Graduate School for research funding. I thank Eric Ravnikar for valuable research assistance. Finally, I gratefully acknowledge financial support from the U.S. Environmental Protection Agency (EPA) under the Science to Achieve Results (STAR) Graduate Fellowship program. EPA has not officially endorsed this publication and the views expressed herein may not reflect the views of the EPA. All errors are my own. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

The Demand for Ethanol as a Gasoline Substitute
Soren T. Anderson
NBER Working Paper No. 16371
September 2010
JEL No. Q41,Q42,Q48

ABSTRACT

This paper estimates household preferences for ethanol as a gasoline substitute. I develop a theoretical model linking the shape of the ethanol demand curve to the distribution of price ratios at which individual households switch fuels. I estimate the model using data from many retail fueling stations. Demand is price-sensitive with a mean elasticity of 2.5–3.5. I find that preferences are heterogeneous with many households willing to pay a premium for ethanol. This reduces the simulated cost of an ethanol content standard, since some households choose ethanol without large subsidies; simulated costs are still high relative to likely environmental benefits.

Soren T. Anderson
Department of Economics
Marshall - Adams Hall, Office 18C
Michigan State University
East Lansing, MI 48824-1038
and NBER
sta@msu.edu

1 Introduction

Policies to reduce oil consumption increasingly promote ethanol and other biofuels through subsidies, mandates, and funding for research. Proponents argue that substituting toward biofuels will enhance energy security, reduce carbon dioxide emissions, improve air and water quality, and benefit farmers. Many recent policies mandate, either explicitly or implicitly, a minimum market share for ethanol. A prime example is the U.S. Renewable Fuels Standard (RFS), which will increase ethanol use to about 25% of gasoline consumption in the coming years. Despite this attention from policymakers, relatively little is known about household preferences for biofuels or the effect that ethanol mandates will have on gasoline markets. This information is critical for designing, implementing, and evaluating policies to promote ethanol and other biofuels.

I address this important research need by estimating demand for ethanol as a gasoline substitute. I find that demand for ethanol is sensitive to relative prices, with an average elasticity of about 2.5–3.5. These are the first available estimates in the literature for the price elasticity of household ethanol demand, which is a key parameter for studies that analyze a retail ethanol subsidy or mandate. I find that elasticities are substantially smaller in magnitude (and less variable) than they would be if fuel-switching behavior were concentrated around a single price. Rather, fuel-switching behavior extends over a wide range of prices where ethanol is discounted 0%–25% below gasoline. These results imply that many households are willing to pay a per-mile premium for ethanol and that preferences for ethanol among these households are actually quite diffuse.

These results have important implications for policy. Previous analyses assume that households are identical and that preferences depend exclusively on ethanol's fuel-economy performance relative to gasoline (Holland, Knittel and Hughes 2008). This assumption can yield misleading results if some households also value ethanol for its perceived environmental and social benefits. In simulations, I find that accounting for households that prefer ethanol can substantially reduce the economic efficiency cost of a hypothetical ethanol content standard (i.e., a minimum market-share requirement), since households with strong preferences choose ethanol without large price subsidies. Similar intuition likely applies for policies to promote other "green" substitutes, such as renewable electricity, energy-efficient light bulbs and appliances, hybrid-electric vehicles, and organic foods.

I begin my analysis by developing a model of household utility in which inputs of ethanol and gasoline combine linearly to produce household transportation services. The key parameter in this model is the relative price at which the household is indifferent between relying entirely on either fuel. When this parameter varies continuously among households, aggregate demand for ethanol is a smooth function of relative fuel prices. Thus, the model formalizes the precise, theoretical link between the distribution of preferences for ethanol and the shape of the aggregate demand curve, allowing me to recover micro preferences from aggregate data.

I estimate the model using a unique dataset that contains nearly 5000 monthly observations for ethanol prices and sales volumes at over 200 individual retail fueling stations in Minnesota during 1997–2006. These data provide a rare and valuable opportunity to document household preferences for biofuels, whose market shares have generally been too small to be included in household surveys or reported separately from gasoline in aggregate measures. I use these data to estimate demand for ethanol as a function of relative fuel prices. Consistent with my theoretical model, which implies that price elasticities might vary dramatically, I estimate demand as a flexible function of relative fuel prices using semi-parametric methods. Previous empirical studies of demand for alternative fuels and gasoline varieties with close substitutes do not allow for this potentially important flexibility.

I use the distribution of preferences implied by my econometric estimates to simulate the effects of a national ethanol content standard. I find that a 25% standard would decrease gasoline consumption by about 20% and would cut carbon dioxide emissions from gasoline by about 10% at an annual efficiency cost of roughly $20 billion. Efficiency losses derive primarily from ethanol's higher marginal production cost. Costs average about $180 per metric ton of carbon dioxide emissions avoided, which is substantially higher than most estimates for marginal external damages, or about $0.80 per gallon of gasoline saved, which exceeds most estimates for the external cost associated with petroleum dependence.

The empirical economic literature on demand for biofuels is miniscule.[1] While an immense literature estimates demand for gasoline, the vast majority of studies focus on the response of overall fuel demand to changes in fuel price levels. Because households have relatively few transportation alternatives, fuel demand in the short run is price inelastic.[2] This paper in contrast focuses on fuel-switching behavior and how demand for ethanol as a gasoline substitute responds to changes in relative fuel prices. Because households that purchase ethanol in my sample are able to substitute easily between ethanol and gasoline, demand for ethanol is price elastic.

Within the fuel demand literature, this paper is most similar to studies that estimate demand for particular fuels with close substitutes, including full-service and self-serve gasoline (Phillips and Schutte 1988) and regular and premium gasoline in both leaded and unleaded varieties (Greene 1989). These studies find own-price and cross-price elasticities that exceed 10 in absolute value. Elasticities also tend to be large for other goods with close substitutes, including breakfast cereals (Nevo 2001), brand-name and generic pharmaceutical products (Ellison, Cockburn, Griliches and Hausman 2006), and individual components of money supply (Barnett, Fisher and Serletis 1992). I improve on this vein of the gasoline demand literature by formalizing fuel-switching behavior in terms of the distribution of household preferences for alternative fuels, using instrumental variables (IV) techniques to identify demand behavior more credibly, and estimating flexible econometric models to test whether elasticities vary with relative fuel prices.[3]

This paper also contributes to the literature showing how to interpret IV estimates in the presence of heterogeneous treatment effects. Following Angrist, Graddy and Imbens (2000), I go

[1] Rask (1998) estimates intermediate demand for ethanol as a 10% blending component in gasoline. He does not estimate household demand. Alves and Bueno (2003) estimate aggregate demand for gasoline in Brazil, which requires 25% ethanol blending in all gasoline, and where ethanol comprises roughly 40% of the non-diesel fuels market (Perkins and Barros 2006). They do not estimate price responses for ethanol. Salvo and Huse (2010) find that the correlation between ethanol and gasoline prices in Brazil increased in the mid 2000s after the introduction of flexible-fuel vehicles, which allow consumers to arbitrage between the two fuels. They do not model heterogeneous preferences, and they do not estimate demand.

[2] Using a variety of methods, Davis and Kilian (2010) estimate price elasticities ranging from roughly -0.1 to -1.1. For relatively recent surveys see Graham and Glaister (2002), Espey (1996; 1998), and Dahl and Sterner (1991). Recent studies indicate that the price response may have declined even further in recent decades (Hughes, Knittel and Sperling 2008; Kilian 2008)

[3] Hausman and Newey (1995) and Yatchew and No (2001) estimate gasoline demand using a semi-parametric approach and other flexible methods. They do not model fuel switching.

beyond the standard testing for whether my instruments predict prices (i.e., F-tests for instrument relevance) and develop a heuristic approach to analyze where in the demand function my instruments actually induce price variation. This analysis allows me to determine which section(s) of the demand function I estimate using instrumental variables. Such an approach may prove useful in future applications.

The format of the paper is as follows. Section 2 discusses the role of ethanol in the fuels market, ethanol's environmental effects, and ethanol production and distribution. Section 3 presents a model of household demand for ethanol as a gasoline substitute, aggregates households to give an expression for market demand, and relates the distribution of household preferences to aggregate price responses. Section 4 describes the data I use to estimate the model, providing descriptive statistics that summarize supply and demand behavior. Section 5 outlines the econometric model, discusses identification, and presents my results. Section 6 uses the distribution of preferences implied by these estimates to simulate the effects of a national ethanol content standard. Section 7 concludes.

2 Background

2.1 Ethanol's role in the fuels market

Ethanol is an alcohol fuel that in the United States derives primarily from corn. Gasoline blenders mix ethanol with gasoline to comply with federal air quality regulations, to produce mid-grade and premium fuels, and to satisfy the federal RFS. Virtually all gasoline vehicles can burn fuel blends that contain 10% ethanol or less. While blenders sometimes use ethanol as a gasoline substitute when ethanol prices are low, ethanol's primary role is as a gasoline complement. Blenders added about 5 billion gallons of ethanol to gasoline in 2006, or about 3.5% of gasoline consumption by volume; blending has since doubled to 10 billion gallons or 7.3% of consumption in 2009. Ethanol is heavily subsidized, with direct federal and state payments to ethanol producers, a federal tax subsidy of $0.45 per gallon for blenders, and a tariff of $0.54 per gallon that applies to all but a nominal quantity of imports.

The market for ethanol as a direct gasoline substitute is small but growing rapidly. Stimulated by rising gasoline prices and supported by federal, state, and local subsidies for alternative-fuel vehicles and infrastructure, the number of retail stations offering E85—an alternative fuel blend of 85% ethanol and 15% gasoline—more than doubled during 2006–2009 to over 1900 stations nationwide. Here and throughout, I refer to E85 simply as "ethanol" or, when necessary to avoid confusion, as "retail ethanol."[4] On the consumer side of this market, the federal Alternative Motor Fuels Act of 1988 created strong incentives under the Corporate Average Fuel Economy (CAFE) standards program for automakers with binding CAFE constraints to produce flexible-fuel vehicles capable of burning both ethanol and gasoline. Automakers produced about 5 million of these vehicles between 2000 and 2006, and production continues apace.

The federal RFS, which Congress first established in 2005 and then expanded in late 2007, sets a minimum quantity of renewable fuel each year from 2008–2022, increasing gradually from 9 to 36 billion gallons. Industry is currently using ethanol to comply with the standard, and this is likely to continue. The quantity standard for 2022 is about 25% of current gasoline consumption. Although the standard mandates a minimum quantity of renewable fuel, the U.S. Environmental Protection Agency (EPA) implements the standard as a percentage of projected fuel consumption. Below I simulate the effects of a 25% ethanol content requirement for gasoline, which is modeled roughly on the RFS for 2022.

Only flexible-fuel vehicles are certified to run on fuel blends containing more than 10% denatured ethanol. These vehicles have larger fuel injectors as well as fuel-system components that are more resistant to corrosion. Earlier models also had special fuel sensors. These components, which increase production costs no more than \$100–\$200, allow the vehicles to burn retail ethanol, regular gasoline, or any combination of the two. Ethanol has lower energy content than gasoline, implying fewer miles per gallon. The ratio of gasoline to retail ethanol mileage is about 1.35,

[4] I distinguish retail ethanol from "denatured ethanol," which is blended with gasoline to produce retail fuels. Denatured ethanol is nearly pure alcohol but with a small quantity of gasoline or other chemical added, making it unfit for human consumption.

which means that retail ethanol's mileage is about $1 - 1/1.35 \approx 25\%$ lower.[5] Thus, households that care only about minimizing fuel costs will demand a 25% price discount for retail ethanol.

2.2 Ethanol's environmental and social effects

It has been estimated that replacing one gallon of gasoline with pure corn-based ethanol reduces net petroleum consumption by 0.95 gallons, after accounting for upstream petroleum inputs and ethanol's lower mileage (Farrell, Plevin, Turner, Jones, O'Hare and Kammen 2006). Ethanol's climate benefits are less impressive. Corn collects energy from the sun and absorbs carbon dioxide from the atmosphere as it grows, but ethanol production from corn is energy-intensive. Corn farming uses a lot of fertilizer, and the ethanol refining process uses a lot of heat. These inputs derive largely from natural gas given current production techniques. As a result, ethanol only reduces net carbon dioxide emissions by 15% after accounting for upstream energy inputs and ethanol's lower mileage (Farrell et al. 2006). In fact, ethanol may in some cases increase emissions, after further accounting for direct and indirect land-use changes associated with growing feedstocks (Searchinger, Heimlich, Houghton, Dong, Elobeid, Fabiosa, Tokgoz, Hayes and Yu 2008; Fargione, Hill, Tilman, Polasky and Hawthorne 2008).

The local air and water quality benefits of ethanol are mixed. Ethanol is an oxygenate that reduces carbon monoxide emissions in older engines, improving air quality, but modern engines and pollution-control equipment largely obviate these benefits. Ethanol reduces tailpipe emissions of benzene (a known human carcinogen) but increases emissions of acetaldehyde (a possible carcinogen) and nitrogen oxide (a precursor to ozone and smog). Ethanol displaces environmentally harmful petroleum refining, but corn production increases fertilizer and pesticide use on environmentally sensitive land. Finally, some policymakers worry about ethanol's role in driving up food prices.

[5] Using Environmental Protection Agency (EPA) estimates for combined city and highway driving, I calculate the ratio of regular gasoline to retail ethanol mileage for each flexible-fuel vehicle model offered between 2000 and 2006 (U.S. Environmental Protection Agency 2000-2006). EPA did not test vehicles using both fuels until 2000, but relatively few flexible-fuel vehicle models were offered prior to 2000. I calculate the sales-weighted mean ratio using data for nationwide sales of individual flexible-fuel vehicle models from the U.S. Department of Transportation.

Household preferences for ethanol as a gasoline substitute vary considerably. First, ethanol's relative mileage varies across vehicles and driving scenarios, even in highly controlled government tests. On the road, some households drive primarily in stop-and-go city traffic, while others log a large fraction of highway miles. These and other differences may affect relative mileage. Second, many households internalize ethanol's perceived benefits. More than half of the drivers in a recent nationwide poll expressed interest in owning a flexible-fuel vehicle (Harris Interactive 2006). Of these, nearly 90% were motivated by reducing oil dependence, while nearly two-thirds wanted to reduce greenhouse gas emissions. Over 90% of the drivers in another poll would prefer to own a flexible-fuel vehicle. When asked about ethanol's benefits, they cited "renewable fuel," "clean fuel," "made in America," and "more economical" with roughly equal frequency (Phoenix Automotive 2006).

2.3 Ethanol production and distribution

There were about 100 ethanol refineries nationwide in 2006, and the number has since doubled to over 200 in 2009. Most refineries are located in the corn belt, although a handful are located outside of the Midwest.

Nearly all denatured ethanol is blended with gasoline in ratios less than 10%. Most blending occurs at fuel blending and distribution terminals, which are located strategically near population centers throughout the country. Terminal operators blend gasoline, ethanol, and other components into finished products and then distribute fuel by tanker truck to individual retail stations. A relatively small share of ethanol blending occurs at ethanol refineries that have infrastructure for fuel blending.

Fuel terminals receive most gasoline by pipeline from oil refineries. Existing pipelines are not suitable for transporting ethanol, however, since ethanol can corrode gasoline pipelines, and since water accumulating in the pipelines can mix with and contaminate ethanol. Moreover, existing pipelines connect large oil refineries with cities, whereas ethanol refineries are usually located in rural areas. In the corn belt, tanker trucks deliver ethanol from ethanol refineries to fuel terminals.

Ethanol traveling from the Midwest to the coasts usually goes by rail.

Retail ethanol is readily available wherever large quantities of denatured ethanol are blended with gasoline. In Minnesota, for instance, retail ethanol is available at virtually every fuel terminal any time of year, because Minnesota has required 10% ethanol blending in all gasoline year-round since October 1997. Terminal operators maintain stocks of fuel and sometimes lease storage facilities to retail chains who manage their own fuel stocks. Retail ethanol is also readily available at a handful of ethanol refineries that have infrastructure for fuel blending. Retail ethanol stations in states such as Minnesota have no difficulty resupplying on short notice, given ethanol's wide availability for gasoline blending.

3 Theoretical model

To motivate my empirical analysis, I develop a model of demand for ethanol as a gasoline substitute. The model formalizes the precise link between the distribution of household preferences and the shape of the aggregate demand function.

3.1 The household's problem

For the moment I assume that each household owns a single flexible-fuel vehicle. The household's utility is quasilinear in transportation services $v(\cdot)$ and other goods:

$$v(e+rg)+x, \tag{1}$$

where $v(\cdot)$ is strictly increasing and strictly concave, e is consumption of ethanol, g is consumption of regular gasoline, x is consumption of all other goods, and r is the rate at which the household converts gallons of regular gasoline into ethanol-equivalent gallons. Ethanol and gasoline are perfect substitutes. That is, utility is defined over a linear combination of ethanol and gasoline, which I call ethanol-equivalent fuel. When a household cares only about miles traveled the conversion rate r exactly equals the ratio of the household's mileage when burning gasoline to its mileage when burning ethanol. This ratio will vary across households due to minor differences in relative mileage. More importantly, some households will value ethanol for its perceived environmental or social benefits.[6] By embodying mileage differences and these other factors, r fully summarizes household preferences for ethanol as a gasoline substitute.

The household's budget constraint is given by

$$y - p_e e - p_g g - x = 0, \tag{2}$$

where p_e and p_g are the prices of ethanol and gasoline, y is income, and I have normalized the price of the composite good to \$1.

Which fuel will the household choose?[7] Because ethanol and gasoline combine linearly in the utility function, the household will be at a corner solution and will purchase ethanol exclusively when $p_e < p_g/r$ and gasoline exclusively when $p_g/r < p_e$. That is, the household will choose the fuel with the lower ethanol-equivalent price. For a household that cares only about mileage, this amounts to choosing the fuel that is least costly per mile. Equivalently, the household will choose ethanol when the conversion rate r is less than the price ratio p_g/p_e. Because the conversion rate r equals the relative price where fuel switching occurs, I also refer to it as the fuel-switching price ratio.

While relative prices determine the type of fuel that a household chooses, quantity demanded depends on absolute price levels, with the household equating the marginal utility of ethanol-equivalent fuel consumption to the ethanol-equivalent price of whichever fuel it chooses. For households that choose ethanol, the optimal quantity of ethanol demanded is therefore given by

$$e^* = q(p_e), \tag{3}$$

[6] In addition, the relative convenience of filling up with ethanol might vary somewhat across households, given the fuel's limited availability; I argue below that this source of variation is not particularly important in my data.

[7] I assume that the household always buys fuel but never spends its full income on fuel. Assuming that $v'(0) > 1$ ensures that the household buys fuel. Assuming that y is sufficiently large, so that $v'(y/p_e) < 1$ and $v'(ry/p_g) < 1$, guarantees that the household does not spend its full income on fuel.

where I have defined ethanol-equivalent fuel demand as $q(p) \equiv v'^{-1}(p)$ given ethanol-equivalent fuel price p. The quantity of gasoline demanded for households that choose gasoline is given by

$$g^* = \frac{q(p_g/r)}{r}, \tag{4}$$

where the presence of r converts ethanol-equivalent gallons into nominal gallons of gasoline. I assume that households that do not own flexible-fuel vehicles (or are otherwise unable to buy ethanol) face the same maximization problem, which implies that their gasoline demand is also given by equation (4).

3.2 Aggregate demand

Because ethanol and gasoline are perfect substitutes, households that own flexible-fuel vehicles sort into ethanol buyers and gasoline buyers according to their fuel-switching price ratios. While each individual household rests at a corner solution, aggregate demand will be a smooth function of relative prices when fuel-switching price ratios are distributed continuously.

To move formally from individual to aggregate demand, I first assume that there are N (technically, an infinite number of measure N) households in the market. Each household owns a single vehicle, and a fraction ϕ of these are flexible-fuel vehicles. I next assume that fuel-switching price ratios are distributed according to the differentiable cdf $H(r)$, defined on $[0,\infty)$. Recall from above that households will choose ethanol if their fuel-switching price ratios are less than the relative price p_g/p_e. So the fraction of households that choose ethanol is simply the cdf evaluated at this relative price: $H(p_g/p_e)$. I assume for convenience that $v(\cdot)$ and flexible-fuel ownership are distributed independently of r (and of each other), so that fuel-switching price ratios are the only relevant source of heterogeneity in the model. I discuss the validity of this independence assumption below.

Given these assumptions, aggregate demand for ethanol as a function of fuel prices is

$$\begin{aligned} Q_e(p_e, p_g) &= N\phi \int_{-\infty}^{p_g/p_e} \bar{q}(p_e) dH(r) \\ &= N\phi H\left(\frac{p_g}{p_e}\right) \bar{q}(p_e). \end{aligned} \tag{5}$$

where $\bar{q}(\cdot) \equiv E[q(\cdot)]$ is expected ethanol-equivalent fuel demand for an individual household (which by independence does not depend on r). Aggregate demand is simply the total number of households, multiplied by the fraction that own flexible-fuel vehicles, multiplied by the fraction of these that choose ethanol (which depends on relative prices), multiplied by average ethanol consumption among households that choose ethanol (which depends on the absolute price of ethanol). The appendix provides similar expressions for aggregate gasoline demand and aggregate welfare, which are important for the policy simulation below.

Taking logs on both sides yields logged aggregate ethanol demand:

$$\ln Q_e(p_e, p_g) = \ln N\phi + \ln H\left(\frac{p_g}{p_e}\right) + \ln \bar{q}(p_e), \tag{6}$$

This equation is critical because it relates fuel prices and ethanol quantities to the distribution of household preferences for ethanol as a gasoline substitute. As is clear from the equation, tracing out the precise shape of the demand curve as a function of relative prices will reveal the underlying cdf of fuel-switching price ratios.

Differentiating (6) with respect to p_g and then multiplying by p_g yields the gasoline-price elasticity of aggregate ethanol demand:

$$\xi_g = \frac{h\left(\frac{p_g}{p_e}\right)}{H\left(\frac{p_g}{p_e}\right)} \frac{p_g}{p_e}, \tag{7}$$

where $h(r) \equiv H'(r)$. This cross-price elasticity quantifies the rate at which consumers switch from regular gasoline to ethanol given a percent increase in the price of gasoline. A 1% increase in gasoline prices leads to a $\xi_g\%$ increase in the quantity of ethanol demanded. Observe that this elasticity is also the elasticity of ethanol's market share (i.e., the fraction of households that choose ethanol) with respect to the price ratio. Thus, I also refer to this elasticity as the fuel-switching elasticity.

Differentiating (6) with respect to p_e and then multiplying by p_e yields the own-price elasticity:

$$\begin{aligned}\xi_e &= \frac{p_e\bar{q}'(p_e)}{\bar{q}(p_e)} - \frac{h\left(\frac{p_g}{p_e}\right)}{H\left(\frac{p_g}{p_e}\right)}\frac{p_g}{p_e} \\ &= \xi_f - \frac{h\left(\frac{p_g}{p_e}\right)}{H\left(\frac{p_g}{p_e}\right)}\frac{p_g}{p_e}, \end{aligned} \tag{8}$$

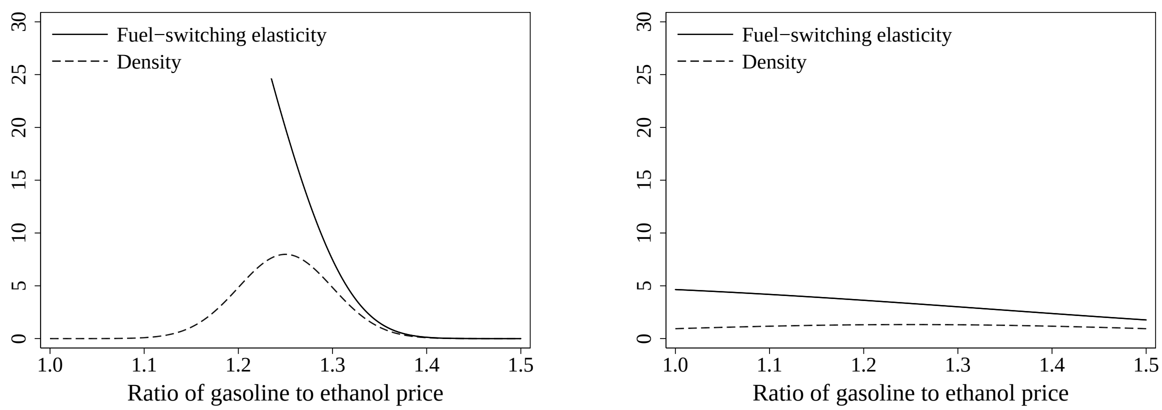
where I have defined $\xi_f \equiv p\bar{q}'(p)/\bar{q}(p)$. The first term in (8), which I refer to as the price elasticity of individual ethanol-equivalent fuel demand, quantifies the rate at which individual households respond to the price increase (on average) by curtailing demand. The second term in (8), which is identical to the gasoline-price elasticity in (7) multiplied by negative one, quantifies the rate at which households switch from ethanol to gasoline as the price of ethanol increases. Again, this is the fuel-switching elasticity (i.e., the elasticity of ethanol's market share with respect to the price ratio), this time multiplied by negative one. Together these terms imply that a 1% increase in ethanol prices leads to a $-\xi_e\%$ decrease in the quantity of ethanol demanded.

As an aside, observe that the fuel-switching elasticity, given by

$$\frac{h\left(\frac{p_g}{p_e}\right)}{H\left(\frac{p_g}{p_e}\right)}, \tag{9}$$

is the hazard rate for exiting the ethanol market as the price ratio decreases. That is, expression (9) gives the instantaneous rate at which households switch to gasoline given a marginal decrease in the price ratio, conditional on choosing ethanol.

Given any distribution of fuel-switching price ratios, equation (7) specifies precisely how elasticities will vary with relative prices. It is clear from the equation that elasticities could vary



(a) $r \sim N(1.25, 0.05)$

(b) $r \sim N(1.25, 0.30)$

Figure 1: Hypothetical preference distributions and elasticity functions

Note: Figure illustrates the relationship between the density function for fuel-switching price ratios and the fuel-switching elasticity, as given by equation (7), for two hypothetical density functions.

dramatically, depending on the shape of the distribution. For this reason, imposing a constant elasticity in empirical applications may yield misleading results. Given a sufficiently flexible estimate of the elasticity function, however, the equation shows how to recover the distribution of fuel-switching price ratios.

Figure 1 illustrates this relationship. When households are nearly identical, as in figure 1(a), fuel-switching behavior is concentrated around a single price ratio, which leads to a large and highly variable price response in that neighborhood. When households are literally identical, as previous studies assume, aggregate demand mirrors individual demand: the entire market is at a corner solution, with all households choosing the fuel with the lowest ethanol-equivalent price. In terms of figure 1(a), this assumption implies a mass point of individuals at the same fuel-switching price ratio, an infinite price response at that single point, and a zero elasticity everywhere else. This extreme assumption has important implications for policy analysis. If ethanol has relatively high costs, so that no ethanol is consumed in the unregulated equilibrium, large distortions in market prices may be required to induce households to choose ethanol.

When households are heterogeneous, however, as in figure 1(b), price elasticities are much smaller in magnitude and less variable. Fuel switching extends over a wide range of relative prices, and demand is not especially sensitive to prices at any particular point. In this case, households with particularly strong preferences for ethanol can be induced to purchase the fuel with less severe distortion of market prices.

In theory, the model also provides a method for disentangling extensive-margin price responses associated with fuel-switching behavior from intensive-margin responses associated with overall fuel demand. Adding equations (7) and (8) demonstrates that the price elasticity of individual ethanol-equivalent fuel demand is simply the sum of the two aggregate elasticities:

$$\xi_f = \xi_e + \xi_g. \tag{10}$$

For a precise quantitative interpretation of this elasticity, consider a simultaneous 1% increase in both fuel prices. No fuel switching occurs, because relative prices do not change, but households that choose ethanol reduce their demand by ξ_f%. Put differently, a 1% increase in the price of ethanol generates both fuel-switching behavior and conservation, while a 1% increase in the price of gasoline only generates the former; thus, the difference in magnitude between these two price responses equals the conservation effect.

4 Data and summary statistics

I estimate the model of logged aggregate ethanol demand in equation (6) above using monthly data for ethanol prices and sales volumes at a large number of retail fueling stations, gasoline prices in those same areas, and several ancillary variables. Table 1 presents summary statistics for my estimation sample.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.
sales volume (gallons)	3352.71	3977.99	6.90	37770.50
retail ethanol price	1.74	0.35	0.74	2.96
retail gasoline price	1.98	0.43	1.10	3.00
retail gasoline / ethanol price	1.14	0.10	0.74	1.69
wholesale ethanol price	1.27	0.56	0.45	3.03
wholesale gasoline price	1.39	0.45	0.44	2.33
wholesale gasoline / ethanol price	1.17	0.33	0.69	2.45
ethanol pump age (months)	29.08	24.27	1.00	110.00
number flexible-fuel vehicles in county	3252.61	4804.87	0.00	24453.00
number ethanol pumps in county	3.72	2.75	1.00	13.00
number gas stations in county	96.69	110.68	4.00	357.00
distance to Benson refinery (miles)	112.47	43.89	4.63	242.18

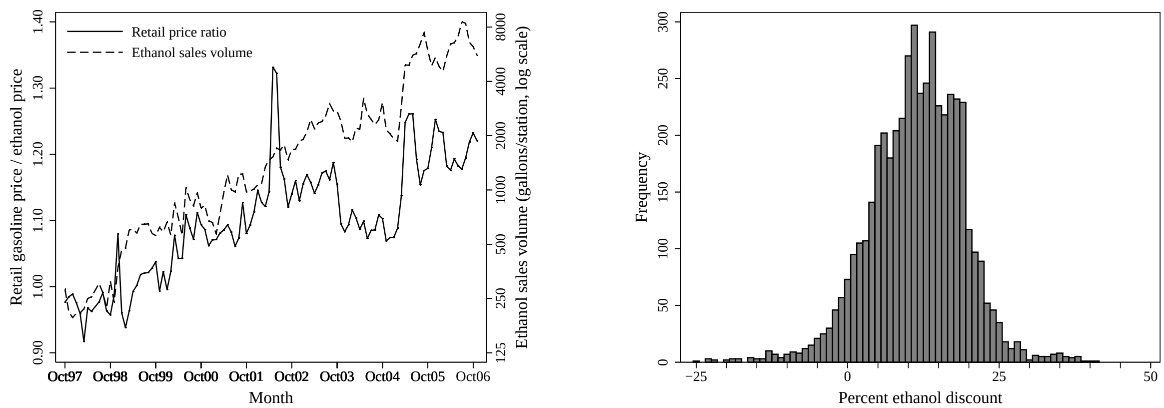
Note: Table is based on estimation sample of 4,825 monthly reports from 232 fueling stations in Minnesota between October 1997 and November 2006. Prices are in 2006 dollars. See text for details.

4.1 Data sources

These data come from several sources. My data for retail ethanol prices and sales volumes come from a Minnesota Department of Commerce and American Lung Association of Minnesota monthly survey of retail ethanol stations in Minnesota. Stations that received funding to help defray ethanol infrastructure costs are required to respond, while other stations may participate on a voluntary basis. This requirement is not strongly enforced, however, and stations do not always report as required. The earliest stations began reporting in October 1997, and the data include records through November 2006.[8]

Stations report volume-weighted prices derived from monthly sales volumes and revenues. Retail prices include federal, state, and local fuel taxes. State and federal fuel taxes did not change during my sample period. The data also record open and close dates for all retail ethanol pumps in Minnesota and the county in which each pump is located. I use this information to calculate

[8] The data include records for a handful of state-operated stations; I ignore these stations in my analysis, because they are only open to government fleets. While government fleets are able to purchase ethanol from private stations, private flexible-fuel vehicles outnumber government flexible-fuel vehicles 100 to 1 in the Midwest Corts (2010). Minnesota's governor issued an executive order requiring state-owned flexible-fuel vehicles to fill up using ethanol "whenever practical," but not until spring of 2006, near the end of my sample period.



(a) Relative prices and sales volumes

(b) Distribution of relative prices

Figure 2: Relative retail prices and ethanol sales volumes

Note: Ethanol sales volume in figure (a) is the monthly average volume of ethanol sales among reporting ethanol stations in Minnesota; the ratio of gasoline to ethanol prices is the volume-weighted sample-mean price of gasoline divided by the volume-weighted sample-mean price of ethanol. Figure (b) is the empirical distribution of relative prices (i.e., percent price discounts for ethanol) in the estimation sample.

the total number of stations operating retail ethanol pumps in each county in each month and the length of time that each pump has been operating, both of which I include as control variables. I match these retail ethanol data to county-average retail prices for regular gasoline from Oil Price Information Service (OPIS). I convert all prices to real 2006 prices using the monthly consumer price index from the U.S. Department of Labor.

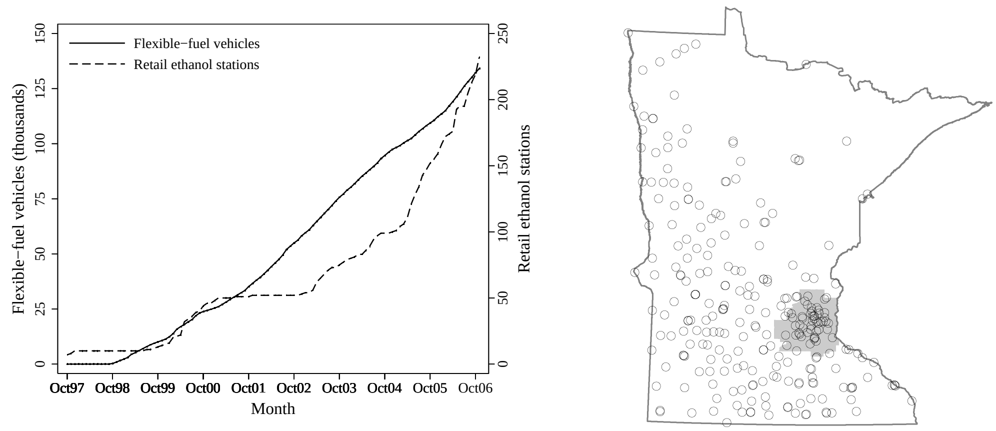
My data report geographic coordinates for many (but not all) stations. Using these coordinates, I attempted to assign brand affiliations (if any) to the stations in my sample.[9] I am unable to identify 14% of stations (accounting for only 5% of my observations), due to missing or inaccurate coordinates, and some of the brand affiliations that I assign to stations are possibly incorrect, due to inaccurate coordinates and changing affiliations over time. While these problems (and station fixed effects) rule out using brand dummies directly in my estimating equation, I do use the brand affiliations to construct my price instruments, as I discuss below.

[9] After locating the coordinates in Google Maps, I searched for the nearest gasoline station and recorded its name. I attempted to corroborate this information using a MN Department of Commerce list of stations operating in Minnesota as of late 2006, a similar list from the National Ethanol Vehicle Coalition, and the U.S. Department of Energy's database of alternative fueling stations (which includes the date the station was added to the database, which is highly correlated with open date).

Figure 2(a) plots relative retail prices over time. Relative prices vary considerably during the sample period, with the relative price of gasoline trending upward. Average ethanol sales also increase steadily over time. The relationship is not necessarily causal, however, as the increase in sales volume is also consistent with a growing stock of flexible-fuel vehicles. I am careful in my estimation to control explicitly for flexible-fuel vehicles and trends in fuel demand. Short-run increases in the relative price of gasoline correlate with contemporaneous increases in ethanol sales volumes, which is perhaps more suggestive of a price response. Again, however, this relationship is not necessarily causal, as unmodeled shifts in aggregate demand might affect fuel prices. Below, I discuss how I identify demand parameters using cross-sectional variation in pricing behavior. Figure 2(b) shows that retailers typically discount ethanol 0%-25% below gasoline; thus, my estimates will reflect price responses within this range of the demand function.

As a measure of underlying fuel costs, I obtain wholesale ethanol price data from a trade publication called *Ethanol and Biodiesel News* (previously known as *Renewable Fuels News* and *Oxy-Fuel News* before that). These data measure weekly spot prices at fuel terminals for denatured ethanol in Minneapolis and Fargo. I assign to each county (and thereby each station) the wholesale price from whichever city is nearest. About four-fifths of stations are located in counties nearest to Minneapolis. I calculate the monthly average of these weekly prices and then subtract the federal ethanol blending subsidy, which fell from \$0.54 per gallon to \$0.51 per gallon during my study period. I obtain wholesale gasoline price data come from the U.S. Energy Information Administration (EIA). These data measure the volume-weighted monthly average spot price in Minnesota. Although wholesale spot price data are available for additional Minnesota cities at a substantial cost from proprietary sources, in practice these prices track each other closely (Minnesota Department of Agriculture 2003). I use these wholesale price variables to construct my price instruments.

In addition to these price variables, I obtain data on flexible-fuel vehicle registrations from the Minnesota Department of Public Safety Division of Driver and Vehicle Services. These data record vehicle identification numbers (VINs), original sales dates, and owner zip codes for all vehicles registered in Minnesota as of the summer of 2007. I identify 154,000 flexible-fuel vehicles in the



(a) Vehicles and stations over time

(b) Spatial distribution of stations

Figure 3: Flexible-fuel vehicles and retail ethanol stations

Note: Figure (a) shows the growth in the number of retail ethanol stations and the stock of flexible-fuel vehicles in Minnesota. Figure (b) shows the locations of Minnesota's 264 retail ethanol fueling stations as of August 2006. Minnesota measures 400 miles from north to south and about 250 miles along its southern border. The shaded region is the seven-county metropolitan area of Minneapolis and St. Paul.

database by cross-referencing VINs with lists of flexible-fuel vehicle models and VIN identifiers from the National Ethanol Vehicle Coalition and from a private firm that collects data on the auto industry. These vehicles represent about 3.3% of the 4.6 million light-duty vehicles registered in Minnesota in 2007. I then use original sales dates to reconstruct a monthly time series for the stock of flexible-fuel vehicles in each county, which I include as a control variable.[10]

Figure 3(a) charts the growth in the number of retail ethanol stations and flexible-fuel vehicles. The flexible-fuel stock grows at a roughly constant rate during the sample period, which is consistent with CAFE standards that generated strong incentives for some manufacturers to produce a limited number of flexible-fuel vehicles each year. Growth in the number of retail ethanol stations accelerated in 2000, when the American Lung Association negotiated an agreement with a particular retail chain to subsidize ethanol pumps at a large number of its stations. Growth accelerated again in 2004-2005. High gasoline prices and low wholesale ethanol costs may have contributed

[10]I am unable to determine whether some vehicles are flexible-fuel vehicles due to missing or invalid VINs, and a relatively small number of flexible-fuel vehicles are excluded due to missing sales dates or zip codes outside Minnesota. I am also unable to account for vehicle attrition or historical movements of vehicles in and out of Minnesota and across county lines prior to 2007. Owner addresses also might differ from counties where flexible-fuel vehicles are actually driven. For these various reasons I measure flexible-fuel stocks with some error.

to this accelerated growth.

As I note above, I calculate the total number of retail ethanol stations in each county in each month to quantify variation in competition, and I include this variable as a control. Figure 3(b) maps the locations for all 264 retail ethanol stations in Minnesota as of August 2006 based on a separate list of station addresses from the Minnesota Department of Commerce. I also calculate the total number of retail gasoline stations operating in each Minnesota county in 2006 based on station address information from the Minnesota Department of Commerce Weights and Measures Division. Table 1, which assumes the same number of gas stations operating in each county for 1997-2006, shows that there are more than 20 gasoline stations for every ethanol station on average in my sample.[11] While competition in fuel markets is typically fierce, most ethanol retailers operate as local monopolists in the narrower retail ethanol market. I use both measures of competition to construct my price instruments.

My analysis covers the time period from October 1997 through November 2006. During this time the number of private retail ethanol stations in Minnesota grew from less than 10 to nearly 250. Based on reported open and close dates, there were about 7500 potential monthly observations at these stations. Approximately 64% of these potential observations are covered by the Minnesota survey. The remaining 36% are missing, reflecting stations that almost never participate in the survey, as well as stations that fail to report in just some months. This results in an estimation sample of 4825 observations at 232 stations, implying an average panel size of about 21 months. Some stations operate nearly the entire study period, while others operate for just a few months, as is clear from figure 3(a).

My data are subject to several potential layers of selection. First, ethanol retailers might locate in areas where preferences for ethanol are strongest. Ethanol pumps are spread throughout Minnesota, however, covering every major region except the sparsely populated northeast, which has

[11]The actual ratio is probably slightly higher. Although most of the nearly 2900 individual stations operating in 2006 were also operating during 1997-2005 (personal communication with Mark Buccelli of the Minnesota Bureau of Weights and Measures), the total number of retail stations statewide declined about 7% from 1997–2006 (National Petroleum News 2006).

higher ethanol transport costs (being farther from ethanol refineries).[12] Rural areas are overrepresented, but infrastructure subsidies in the state were allocated so as to make ethanol as widely available as possible. Minnesota itself has more pumps than other states, but the state has lower ethanol transport costs (having many in-state ethanol refineries) and has been receiving federally funded infrastructure subsidies longer than most states.

Second, not all stations participate in the Minnesota survey, not all participating stations report every month, and stations appear and disappear from the sample as they open and close over time. Below, I test formally for biases related to an unbalanced panel and sample selection in my dataset, finding no evidence for either.

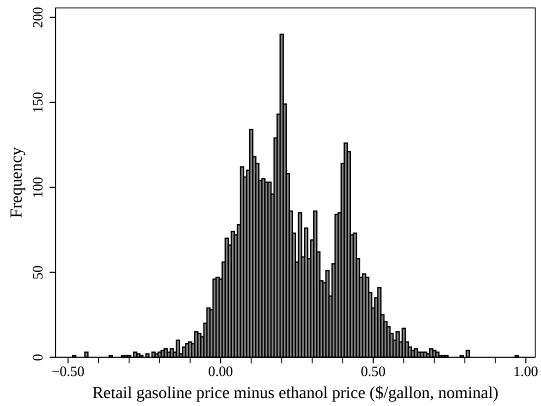
Finally, flexible-fuel owners might have systematically different preferences than other drivers. Flexible-fuel owners tend to buy American, and they are more likely than other drivers to consider minivans and pickups for their next purchase (Phoenix Automotive 2006). Furthermore, most flexible-fuel vehicles have identical gasoline-only counterparts, which could lead to sorting directly on flexible-fuel capacity. On the other hand, automakers produce flexible-fuel vehicles primarily to comply with fuel-economy regulations. They market the vehicles all over the country, even in areas where ethanol is not available, they sell the vehicles for the same prices as comparable gasoline-only vehicles, and flexible-fuel buyers are not observably different from buyers of comparable gasoline vehicles (Anderson and Sallee 2010). In sum, while there is undoubtedly some selection present in my data, it is not necessarily severe. Consistent with this judgment, I find below that price responses do not vary significantly across sub-samples.

4.2 Retail pricing behavior and instruments

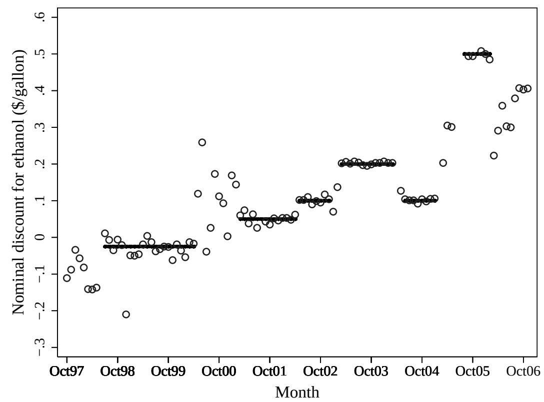
I spoke with industry representatives and inspected retail pricing behavior closely to identify price variation that is arguably exogenous to demand.[13] Retailers generally set prices using rule-of-

[12] Corts (2010) finds that ethanol availability within the Midwest is highly correlated with the presence of flexible-fuel vehicles and proximity to ethanol refineries.

[13] I spoke with representatives from the largest retail chains in Minnesota that offer retail ethanol, as well as several independently owned and operated stations, representatives from two ethanol refineries that directly supply about one-third of retail ethanol stations in Minnesota, several ethanol industry analysts, and the administrators of the Minnesota



(a) Distribution of discounts



(b) Discount over time (example)

Figure 4: Nominal price discounts

Note: Figure (a) shows the empirical distribution of ethanol's nominal retail price discount relative to gasoline in the estimation sample. Figure (b) shows ethanol's nominal price discount relative to gasoline for an example ethanol retailer over time.

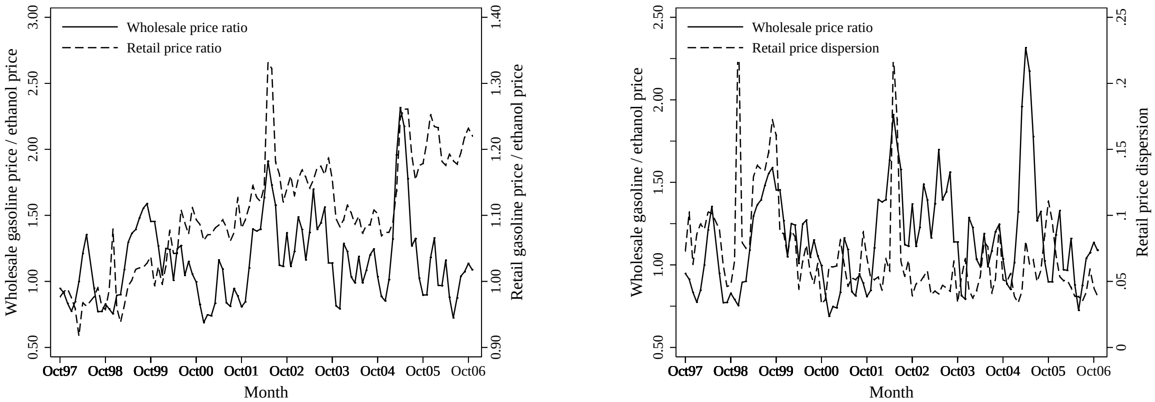
thumb strategies. Most retailers price ethanol at a discount to regular gasoline in nominal increments of $0.10 per gallon, while some price at a fixed nominal markup over wholesale ethanol. This behavior is manifest in figure 4(a), which plots the distribution of nominal price discounts in my sample, and is consistent with how retailers tell me they set prices.[14]

The industry representatives with whom I spoke indicated that discounts at individual stations can sometimes persist at the same level for extended periods. Retailers update discounts primarily to adjust for broad shifts in relative fuel costs; they do not deliberately adjust discounts in response to local, short-term shifts in demand, which are probably not even detectable until after the fact. This behavior is evident in figure 4(b), which plots the nominal discount for one station over time. This station has been operating longer than most but its pricing behavior is fairly typical. This behavior helps identify demand parameters: because retailers often set rule-of-thumb discounts and maintain them for extended periods of time, unmodeled shifts in ethanol demand will tend not to affect relative fuel prices, limiting the potential for price endogeneity.

The sizes of these discounts depend on underlying fuel costs as determined in the broader

survey.

[14] While this pricing behavior may appear suboptimal, ethanol sales are very low relative to gasoline (and not hugely sensitive to prices, as I show below), so the stakes are not especially high.



(a) Relative wholesale and retail prices

(b) Dispersion of relative retail prices

Figure 5: Relative wholesale prices and relative retail fuel prices

Note: The ratios of gasoline to ethanol prices in figures (a) and (b) are the volume-weighted sample-mean prices of gasoline divided by the volume-weighted sample-mean prices of ethanol. The dispersion of the retail price ratio in figure (b) is the monthly standard deviation of the OLS residuals from the retail price ratio regressed on a vector of month and station dummies. This variable quantifies differential changes in relative prices across stations.

markets for gasoline and fuel additives. Average discounts generally increase when wholesale prices for denatured ethanol fall relative to gasoline, and discounts shrink when ethanol prices rise, as is evident in figure 5(a). The economic causality is markedly one sided: events specific to the tiny retail ethanol market have zero bearing on prices for crude oil, gasoline, or even denatured ethanol, whose primary role is as a fuel additive. See the appendix for an extended discussion on the determinants of wholesale fuel prices.

The key to my identification strategy is that, for a variety of reasons, these changes in market spot prices affect the ethanol retailers in my sample differently. One reason is the different relationships the retailers have with their suppliers. As of 2006, about one third of ethanol retailers in Minnesota bought finished fuel from an ethanol refinery in Benson, which is a small town in the southwestern part of the state. Throughout the entire sample period, this refinery supplied ethanol to retailers at a fixed nominal discount below the spot price of gasoline. The retailers, in turn, agreed by contract to pass this same discount along to consumers at their stations. [15] As a

[15] This unique pricing agreement ended in the fall of 2007. The ethanol refinery now supplies retail ethanol at market prices, and retailers are free to price ethanol at whatever price the market will bear. I am not aware of any

result, relative prices are mechanically less variable at these stations. Other retailers have no such contractual arrangement and therefore bear the full brunt of variation in relative spot prices, which transmits to the retail level. Even among these other retailers, variation in contracts could lead to variation in pricing behavior.[16]

Differences in local competition will also lead to variation in pricing behavior. Retail ethanol prices will be more sensitive to changes in relative fuel costs for retailers facing greater competition from other ethanol retailers, whereas retailers in less competitive areas will price largely based on willingness to pay and will therefore be less sensitive to costs. In addition, stations facing different levels of competition in the overall fuel market will also price differently. Where competition is weak, consumers that do not buy ethanol at a given station will likely buy gasoline from the same station instead; where competition is fierce, these consumers are more likely to buy gasoline from a different station altogether. Thus, differences in overall competition will also lead to variation in pricing behavior. I make these points formally in the appendix.

This variation in pricing behavior is critical: it allows me to control for month effects common to all stations and still identify demand parameters using differential changes in fuel prices across stations. To quantify this variation, I regressed relative prices on a vector of month and station effects and then computed the standard deviation of the residuals in each month. I refer to this standard deviation as the dispersion of relative prices. Figure 5(b) shows that price dispersion increases when gasoline spot prices are high relative to ethanol. This behavior is consistent with the different supply relationships I document. Some stations have supply contracts that mechanically tie retail ethanol prices to gasoline, while other stations purchase ethanol at spot prices and pass these costs on to consumers. Price dispersion therefore increases whenever ethanol and gasoline spot prices diverge. This behavior is also consistent with differences in competition. When ethanol

---

similar agreements between ethanol retailers and their suppliers.

[16]Larger retail chains generally have long-term contracts for denatured ethanol and blend their own fuel, whereas smaller chains and independents buy fuel from terminal operators at market spot prices. Contract prices for denatured ethanol are often tied directly to the price of gasoline, which means that relative fuel costs are less variable for larger firms. In theory, the opportunity to sell fuel on the spot market should equalize marginal costs across these firms. Only a small fraction of denatured ethanol actually trades on the spot market, however, and so it is possible that larger firms with long-term contracts perceive their ethanol costs as being less variable. If so, then ethanol's relative price may be less variable at their retail stations.

costs are low relative to gasoline, competitive retailers are forced to reduce prices, while less competitive retailers are able to price closer to gasoline. Price dispersion therefore decreases as the gap between ethanol costs and gasoline prices narrows.

To exploit this variation in pricing behavior, I construct three distinct sets of instruments. The first set (my "Brand" instruments) interact logged wholesale prices for ethanol and gasoline with dummy variables for the 14 identifiable retail brands in my sample (28 variables total). These instruments predict variation in pricing behavior related to chain-specific supply relationship and idiosyncratic pricing strategies. The second set (my "Benson" instruments) interact logged wholesale prices with the logged distance between each county's population-weighted center and the Benson ethanol refinery (2 variables total). These instruments predict variation in pricing behavior related to having a supply contract with the Benson refinery. Because ethanol is costly to transport, the Benson refinery is most likely to supply nearby stations. The third set (my "Competition" instruments) interact logged wholesale prices with the logged numbers of ethanol and gasoline retailers in each county (4 variables total). These instruments help predict variation in pricing behavior related to differences in local competition. I use these three sets of instruments to identify variation in relative prices that is arguably exogenous to demand.

5 Econometric estimation and results

5.1 Econometric model

I estimate logged aggregate ethanol demand of the following form:

$$\ln volume_{it} = \alpha \ln p_{e_{it}} + F\left(\ln \frac{p_{g_{it}}}{p_{e_{it}}}\right) + \beta' X_{it} + \gamma_t + \delta_i + \omega_i(t) + \varepsilon_{it}, \tag{11}$$

where: $volume_{it}$ is gallons of ethanol sold at fueling station i in month t; $p_{e_{it}}$ is the retail price of ethanol and $p_{g_{it}}$ is the retail price of regular gasoline; X_{it} is a vector of time-varying county and station characteristics; γ_t is a month effect that is constant across all fueling stations; δ_i is a fueling station effect that is constant across all time periods; $\omega_i(t)$ is a station-specific time trend; ε_{it} is an unobserved station-month demand shifter; and the remaining elements are coefficients, vectors of coefficients, and functions to be estimated. Note that regression (11) is the empirical analog of logged aggregate demand in theoretical equation (6) above.[17]

While my main estimates assume that $F(\cdot)$ is linear, implying a constant fuel-switching elasticity, my theoretical model implies that elasticities may vary dramatically with relative prices. I test for variable elasticities using two approaches. First, I estimate the model using different sets of instruments. In the presence of a variable elasticity response, different instruments may yield different estimates, if those instruments are inducing price variation at different points in the demand function Angrist et al. (2000). I return to this issue below when interpreting my main results. Second, I use OLS to estimate flexible polynomial, cubic spline, and semi-parametric approximations for $F(\cdot)$. While it is well-known that OLS is a biased estimator of demand, I argue below that the OLS bias in my application relatively mild.

The own-price elasticity of ethanol demand in this model is simply $\alpha - F'(\ln p_g/p_e)$. The gasoline price elasticity is $F'(\ln p_g/p_e)$, which is equivalent to the fuel-switching elasticity. Following equation (10), the price elasticity of individual ethanol-equivalent fuel demand is the sum of the gasoline-price and own-price elasticities (i.e., the difference in magnitudes), which simplifies here to α. Thus, equation (11) imposes a constant price elasticity for overall fuel demand, which is consistent with recent nonparametric estimates (Yatchew and No 2001). When $F(\cdot)$ is linear, this model is equivalent to the standard linear-in-logs demand model, with constant own-price and cross-price elasticities.

In my main estimates, I impose $\alpha = -0.20$ rather than estimate it directly; this value is consistent with previous estimates for the short-run price elasticity of fuel demand. I do this for two reasons. First, efficiency. To pin down α precisely, I need to observe different stations charging the

[17] Unfortunately, I do not observe gasoline quantities below the state level, and so I am unable to calculate local market shares. I am also unable to estimate the model using any alternative level of aggregation (e.g., zip code or county), because I do not observe prices and sales volumes for every ethanol station. Ethanol stations tend to be isolated from one another, however, with stations in my sample located 8 miles from their nearest competitors on average. Thus, it is valid to treat the stations themselves as approximately distinct ethanol markets.

same relative prices (so that fuel choice is held constant) but different price levels (so that quantity demanded varies). Unfortunately, because gasoline prices vary little after controlling for month and station effects, the variation needed to pin down α precisely does not exist, and attempting to estimate it inflates the standard errors on the fuel-switching responses.[18] Second, consistency. While I argue that most variation in relative prices is orthogonal to demand, this same argument does not hold for price levels. Thus, imposing α helps me mitigate endogeneity problems when I estimate the model using OLS. I show below that my fuel-switching results are not particularly sensitive to the choice of α, whether imposed or estimated freely.

Returning to the econometric model, the fueling station effect δ_i controls for persistent differences in fueling station characteristics, such as brand name, location, and amenities. The station effect also controls for persistent determinants of local fuel demand, including household income and other demographics, driving habits, and vehicle efficiency. The month dummy variables given by γ_t control for trends in demand related to growing awareness of flexible-fuel vehicle capabilities or rising state income levels. The station-specific time trends $\omega_i(t)$ control for similar factors that evolve at different rates locally. Finally, the month dummies control for potential seasonality in demand, including the well-known surge in driving that occurs each summer.[19]

The vector of time-varying station characteristics X_{it} includes the log of the county's flexible-fuel vehicle stock. The vector also includes the log of the total number of stations that offer retail ethanol in the same county. While a negative coefficient would imply that new stations draw customers away from existing stations, a zero coefficient might only suggest that new stations locate where competition is weak. This measure of competition reflects retailer choices about when and where to install ethanol pumps, and these decisions presumably depend critically on the locations of existing pumps. Table 1 indicates that there are less than 5 retail ethanol stations per

[18] After controlling for station effects, the month dummies explain just 42% of the variation in logged relative prices. The remaining variation comes almost entirely from differences in ethanol prices across stations: the month dummies explain 99.5% of the variation in logged gasoline prices but only 88% of the variation in logged ethanol prices.

[19] The minimum denatured ethanol content of retail ethanol in Minnesota varies seasonally due to cold weather starting issues, ranging from 70% in the winter to 79% in the summer (U.S. Department of Energy 2006). Although the month dummies control for seasonality in the level of demand, they do not control for potential seasonality in the price elasticity of demand due to variation in denatured ethanol content. Variation in ethanol content is relatively minor, however, and unlikely to be transparent to consumers, making it neither problematic nor useful for identification.

county, while there are more than twenty times as many gasoline stations. A finding of significant competition in retail ethanol markets would therefore be surprising. Finally, the vector of time-varying station characteristics includes dummy variables that indicate the length of time that a station has been offering ethanol. These dummy variables differ from the month dummy variables because start dates vary from station to station. Sales will likely be low after a station first opens before customers are fully aware of the new opportunity to purchase ethanol.

5.2 Identification

I estimate regression (11) using ordinary least squares (OLS) and two-stage least squares (2SLS). OLS estimates are potentially biased if unmodeled shifts in ethanol demand correlate with fuel prices. This is a standard endogeneity problem in estimating demand functions. Shifts in ethanol-specific demand would tend to bias the own-price elasticity toward zero, if such shifts led to higher ethanol prices. In contrast, shifts in overall fuel demand would tend to increase prices for all fuels, in which case relative prices would arguably be exogenous. This would facilitate identification using OLS because I am primarily interested in fuel-switching behavior, which only depends on relative prices. Endogenous price levels would nevertheless bias OLS estimates for the price elasticity of individual ethanol-equivalent fuel demand (when this parameter is estimated freely).

In practice, the station owners with whom I spoke indicated that they do not deliberately update retail ethanol prices in response to local, short-term demand shifts. Rather, they price ethanol at nominal discounts to regular gasoline (or markups over denatured ethanol), often maintain these discounts for extended periods of time, and only adjust discounts in response to changes in underlying fuel costs. Indeed, for a large fraction of stations in the sample, discounts are fixed by contract. This behavior largely rules out ethanol-specific demand shifts at individual stations being correlated with station-level price changes and biasing OLS estimates.[20]

[20] There is theoretical justification for retailers being unresponsive to local demand shifts when setting relative prices. For a monopolist ethanol retailer, relative retail prices will be invariant to demand shifts that enter multiplicatively by scaling aggregate demand. This is because multiplicative demand shifts do not alter the shape of the own-price elasticity function, leaving the monopolist's first-order pricing condition unchanged. See the appendix for benchmark models of retail pricing behavior.

Underlying fuel costs could still be endogeneous to local demand shifts, however, if such shifts were correlated across many stations. That is, even if individual retailers are price takers in wholesale markets, their collective behavior could influence wholesale prices, meaning that wholesale prices are not exogenous in an econometric sense (Kennan 1989). A classic example is the surge in travel demand that drives up fuel prices each summer. I control for these and other correlated demand shifts using month dummy variables. Finally, I control for any slowly evolving local demand shifts using station-specific trends.

While these controls throw away potentially useful time-series and cross-sectional variation, I am able to document a variety of contractual relationships between retail ethanol stations and their wholesale suppliers, as well as variation in local competition, which lead to cross-sectional variation in pricing behavior. To exploit this variation in pricing behavior, I construct three sets of price instruments as described above. These instruments included logged wholesale ethanol and gasoline prices interacted with: (1) station brand dummies, (2) logged distance to the Benson refinery, and (3) the logged numbers of ethanol and gasoline stations operating in the same county. In effect, I am treating as exogenous the price variation that derives from different rule-of-thumb pricing strategies, different supply relationships, and different levels of market competition interacting with wholesale fuel prices, even though wholesale prices themselves are not necessarily exogenous. In my OLS estimates, I also retain variation related to the idiosyncratic timing of when individual stations adjust their rule-of-thumb discounts; this variation is valid so long as the size and timing of such adjustments is exogenous, conditional on month and station effects.

Additional identification issues arise in the context of an unbalanced and non-random sample of stations. Stations appear and disappear from my dataset as they open and close pumps, join the Minnesota survey, or fail to report. As long as these choices are uncorrelated with demand, conditional on controls, then OLS estimates are consistent (Wooldridge 1995 2002). This seems plausible, given that I explicitly control for fuel prices, station effects, month dummies, and other likely determinants of selection. I tested this hypothesis formally by adding leads and lags of selection indicators to the regression in equation (11). The F-statistic on these variables was highly insignificant, suggesting that standard selection bias is not a problem.[21]

A separate but related issue is that stations with long panel lengths will weigh heavily in the estimates relative to stations with short panels, while stations without ethanol pumps receive no weight at all. This is not a concern if the elasticity function $F'(\cdot)$ is the same everywhere. If the elasticity function varies over time or across stations, however, and if the stations I observe are not representative, then my estimates of the "average" elasticity function will be biased. I examine this issue below by estimating price responses separately for different time periods and for different regions. I find no significant disparities.

5.3 Estimation results

5.3.1 Constant elasticity estimates

Table 2 presents my main OLS and 2SLS estimation results, which impose $\alpha = -0.20$ and also assume a constant price elasticity for fuel-switching behavior. I control for station effects using both fixed-effects and first-difference estimators, which have different efficiency properties in the presence of serial correlation and different probability limits in the presence of dynamic price responses. Below, I test the sensitivity of the results to alternative values of α and relax the assumption of a constant fuel-switching elasticity.

Ethanol demand is sensitive to price changes. The coefficient on logged relative prices in regression (1), which is based on the OLS fixed-effects estimator, implies that the elasticity of ethanol's market share with respect to relative prices is 2.730. The same coefficient is 3.484 in equation (2), which is based on the 2SLS fixed-effects estimator. These results imply that the OLS estimator is biased toward zero, which is consistent with the usual intuition. The implied bias is only about 20%, however, which is consistent with my arguments above that most price variation is orthogonal to demand. The corresponding elasticities based on the first-difference estimator, in regressions (3) and (4), are about 0.9 smaller in magnitude. Why? One possible explanation

[21] I added one-period leads and lags of a dummy variable, call it s_{it}, that equals one if I observe data for station i in month t and zero otherwise. I also added $\sum_{r>t}^{T} s_{ir}$. Wooldridge (1995; 2002) suggests adding $\sum_{r\neq t}^{T} s_{ir}$ and $\prod_{r\neq t}^{T} s_{ir}$, but neither has any time variation in my panel.

Table 2: Main estimation results

	Fixed effects		First differences	
Variable	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
ln(gas price / ethanol price)	2.730	3.484	1.872	2.613
	(0.193)	(0.406)	(0.233)	(0.562)
ln(number flex-fuel vehicles)	0.048	0.059	0.095	0.096
	(0.023)	(0.024)	(0.023)	(0.023)
ln(number ethanol stations)	-0.097	-0.092	-0.035	-0.013
	(0.077)	(0.076)	(0.092)	(0.094)
month 1 of operation	-0.648	-0.636	-0.688	-0.690
	(0.100)	(.099)	(0.112)	(0.112)
month 2 of operation	-0.075	-0.069	-0.105	-0.106
	(0.078)	(0.077)	(0.083)	(0.083)
month 3 of operation	0.003	0.014	-0.033	-0.032
	(0.068)	(0.068)	(0.058)	(0.058)
month 4 of operation	-0.014	-0.007	-0.036	0.034
	(0.046)	(0.046)	(0.032)	(0.033)
Number of observations	4825	4825	4148	4148
Number of stations	232	232	202	202
R-squared	0.18		0.13	
Residuals AR(1)	0.422	0.444	-0.249	-0.246
	(0.029)	(0.027)	(0.026)	(0.026)
F-statistic (weak instruments)		18.83		68.61
(Chi-square p-value)		(0.00)		(0.00)
Hansen's J-statistic (overidentification)		45.956		34.165
(Chi-square p-value)		(0.066)		(0.412)

Note: Dependent variable is logged monthly ethanol sales volume in gallons; results impose an overall fuel demand elasticity of -0.20. Clustered standard errors (in parentheses) are robust to arbitrary heteroskedasticity and serial correlation within stations. All regressions control for station effects, month dummy variables, and station-specific time trends; R-squared is the fraction of remaining variation explained by the variables above. Residuals AR(1) is the coefficient from the least-squares regression of the residuals on their lagged values. F-statistic (for weak instruments) tests the null that excluded instruments have no explanatory power in the first-stage regression; robust p-values are in parentheses. Hansen's J-statistic (for overidentification) tests the null that instruments are jointly uncorrelated with the errors; robust p-values are in parentheses. See text for details.

is that demand does not respond fully to changes in relative fuel prices within the first month, in which case the fixed-effects and first-difference estimators may give different results. The first-difference estimator exploits the correlation between price and quantity changes in adjacent time periods only, while the fixed-effects estimator relates average sales volumes to relative fuel prices in all time periods. For this reason, fixed-effects estimates may be more robust to delayed price responses.[22]

[22]Indeed, when I include lagged price variables, the fixed-effects and first-difference OLS estimates begin to con-

The instruments appear to be performing well. The first-stage F-statistics are highly significant: they indicate that the instruments are strong predictors of relative fuel prices, conditional on covariates. At the same time, the Hansen's J-statistics for overidentification are not significant: I am unable to reject the null that the instruments are jointly uncorrelated with the error terms in the model (and that the price elasticity is constant, as I discuss below), although the statistic is borderline significant in the fixed-effects 2SLS model.

Standard errors in table 2 are robust to arbitrary heteroskedasticity and serial correlation. The fixed-effects estimates have slightly narrower confidence intervals than the first-difference estimates. The errors are serially correlated for both estimators, however, and neither estimator is fully efficient.[23]

Table 3 presents elasticity estimates when imposing different values of α (in the first four rows) and estimating α freely (in the last row). The table omits results for the covariates, since they change little. A clear pattern emerges: as the imposed value of α increases in magnitude, the fuel-switching elasticity decreases by roughly the same magnitude. This is not surprising: logged ethanol prices and logged relative prices are almost perfectly collinear; hence, the offsetting effects. When I attempt to estimate α freely, the standard error on the fuel-switching effect increases substantially, while α itself is estimated imprecisely and has the wrong sign.[24] I do not put much stock in these estimates, however, given the inherent difficulty in pinning down α in this model. In

verge. Another possible explanation is I calculate relative fuel prices based on county-average gasoline prices. While it is unclear that a different level of aggregation is more appropriate, measurement error will tend to bias the elasticity estimates toward zero, and this bias is usually more severe in first-difference estimates (Griliches and Hausman 1986). The 2SLS estimator should correct for this bias, however, and so I suspect this is not the primary explanation.

[23]First-order serial correlation in the fixed-effects residuals is about 0.40 and statistically different from zero. First-order serial correlation in the first-difference residuals is -0.25. This coefficient is statistically different from zero, which indicates that the first-difference estimates are not efficient. This coefficient is also statistically different from -0.5, which confirms the inference based on the fixed-effects residuals that the model's errors in levels are serially correlated (Wooldridge 2002). That $-0.25 \approx -(1-0.40)/2$ is consistent with the model's errors following an AR(1) process (Solon 1984).

[24]There are several possible explanations. First, a household's overall fuel demand may be correlated with its fuel-switching price ratio, which would violate my assumption that they are independent. Second, some households may be responding to linear differences in fuel prices instead of relative prices. When I add the linear difference to the models estimated using OLS fixed-effects and first-differences, however, its coefficient is insignificant, while α continues to have the wrong sign and be insignificant. Lastly, if consumers respond to price changes with a delay, and this delay is longer for ethanol prices, this could manifest as a positive coefficient on α. When I add lagged price effects, the coefficient on α flips signs for fixed-effects estimation, but remains highly insignificant.

Table 3: Sensitivity to different choices for α

Value of α	Fixed effects OLS	Fixed effects 2SLS	First differences OLS	First differences 2SLS
$\alpha = -0.00$	2.927	3.679	2.069	2.816
	(0.193)	(0.406)	(0.232)	(0.561)
$\alpha = -0.10$	2.829	3.582	1.970	2.715
	(0.193)	(0.406)	(0.232)	(0.562)
$\alpha = -0.20$ (main results)	2.730	3.484	1.872	2.613
	(0.193)	(0.406)	(0.233)	(0.562)
$\alpha = -0.30$	2.632	3.387	1.774	2.511
	(0.194)	(0.406)	(0.233)	(0.563)
$\alpha =$ freely estimated	3.465	4.320	2.777	3.613
	(0.616)	(3.669)	(0.517)	(1.224)
α itself	0.545	0.656	0.720	0.782
	(0.607)	(3.739)	(0.433)	(1.235)

Note: Table replicates the results in table 2 above while imposing different values for α (in the first four rows) and estimating α freely (in the last row). Clustered standard errors (in parentheses) are robust to arbitrary heteroskedasticity and serial correlation within stations. See text and the previous table for details.

any case, the fuel-switching response is relatively stable, and I am unable to reject $\alpha = -0.20$ or any of the other reasonable values I impose in table 3.

Returning to the estimates in table 2, the coefficients on flexible-fuel vehicle stocks indicate that a 1% increase in the number of vehicles leads to a 0.05%-0.10% increase in ethanol sales volumes. I had expected to find coefficients closer to 1, indicating that ethanol sales increase proportionally with the density of potential buyers. I suspect that this estimate is biased toward zero, however, as a result of measurement error, which is exacerbated in panel data models (Hausman 2001).[25]

The coefficients in the next row indicate that a 1% increase in the number of ethanol pumps per county leads to a 0.01%-0.10% reduction in sales volumes at individual stations; these coefficients are not statistically different from zero. Conditional on where retailers choose to locate, new pumps draw only a small fraction of customers away from existing stations. This result is not surprising,

[25] Using my monthly panel of flexible-fuel stocks, I regressed the logged number of flexible-fuel vehicles on a vector of station and month dummy variables. These controls explained 88% of the variation in flexible-fuel stocks. Any residual variation that remains is likely contaminated by measurement error, given that I construct my panel using a snapshot of vehicles on the road in 2007. In addition to being noisy, my measure of flexible-fuel stocks is likely biased, as I systematically undercount vehicles from earlier time periods that may have exited Minnesota or been scrapped prior to 2007. This is not a problem if the rate of exit and scrappage is similar across counties, however, because I include month dummies.

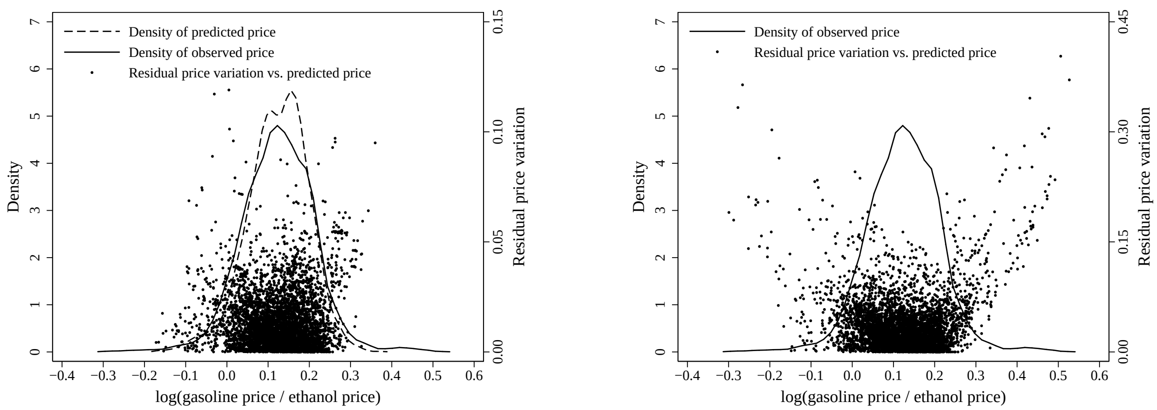
given the small number of ethanol pumps statewide and the fact that infrastructure subsides targeted areas where ethanol was not already available.

The last set of coefficients indicate that sales volumes are low in the first several months after a pump begins operating. Sales volumes are about $\exp(-0.70) - 1 \approx 50\%$ lower in the first month but quickly increase to long-run levels within a month or two. This rapid increase indicates that market participants are well-informed about ethanol's availability.

5.3.2 Instrumental variables interpretation

Angrist et al. (2000) show how to interpret linear (in logs) IV estimates of demand when the underlying demand function is nonlinear. They show that IV returns a weighted average of the elasticity response over the entire demand function. Roughly, weights are proportional to the density of the data within the range of prices over which the instruments induce price variation. So, for example, if the instruments only generate price variation in the high-price region of the demand function, IV will reflect price elasticities in that region only. This interpretation is analogous to the local-average treatment effect (LATE) interpretation for IV in the case of a discrete-valued treatment variable.

This result has two implications. First, OLS and 2SLS estimates may differ, even when OLS is unbiased, if the estimators are implicitly estimating different sections of the demand function. To explore this possibility, I calculated the density of predicted prices from my first-stage regression. Figure 4(a) shows that the density of predicted prices overlaps closely with the density of observed prices. Next, I calculated for each observation the marginal contribution that the instruments make toward predicting prices in the first-stage regression (i.e., the absolute value of the inner product of the instruments with their first-stage coefficients). Figure 6(a) shows that the instruments induce price variation everywhere, while figure 6(b) shows that the analogous identifying variation in OLS is similarly distributed (except in the extremes of the data for a handful of observations). Thus, I conclude that the OLS and 2SLS estimators are applying roughly similar weights to different sections of the demand function.



(a) Identifying variation in 2SLS

(b) Identifying variation in OLS

Figure 6: Identifying variation in 2SLS and OLS estimates

Note: Figures show the identifying variation in 2SLS and OLS. Figure (a) shows: (1) the distribution of first-stage predicted prices from 2SLS (as well as the distribution of observed prices for comparison) and (2) the absolute value of the inner product of the instruments with their first-stage coefficients scatter-plotted versus the first-stage predicted prices themselves. Figure (b) shows the analogous information for OLS: (1) the distribution of observed prices and (2) the absolute value of the residuals from a regression of observed prices on covariates scatter-plotted versus observed prices themselves. Note that 100 times logged relative price is approximately equal to ethanol's percent discount. See text for details.

A second, related implication is that different sets of instruments may give statistically different estimates, either because one or more of the instruments is endogenous (the usual interpretation) or because the instruments are estimating different sections of the demand function (the LATE-type interpretation). I explore this issue by estimating model (11) separately using different combinations of instruments. Table 4 presents the estimated fuel-switching price responses from these 2SLS regressions, as well the first-stage F-tests (testing weak instruments) and Hansen's J-statistics (testing overidentifying restrictions). The F-statistics are all significant, implying that each subset of instruments does a good job of predicting prices. Looking down the columns of fixed-effects and first-difference 2SLS coefficients, there is a noticeable pattern: models that include the brand instruments have larger estimated elasticities. In general, however, the estimates are fairly consistent across the different sets of instruments, and the J-statistics (with two exceptions) are all insignificant. I take this as indirect evidence that the elasticity function is roughly constant over the range of observed prices; I explore this issue in further detail below.

Table 4: Robustness to alternative sets of instruments

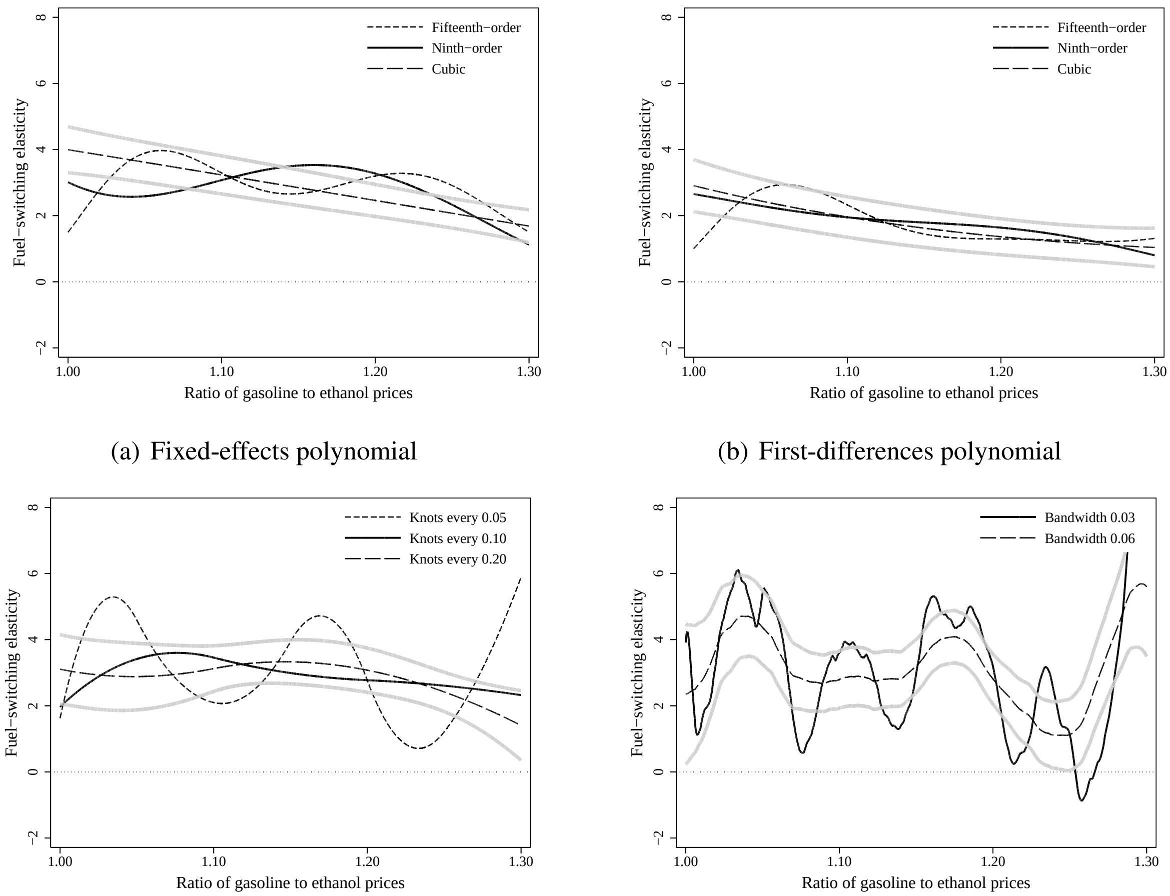
	Fixed effects			First differences		
Instrument set	Coefficient	F-statistic	J-statistic	Coefficient	F-statistic	J-statistic
None (main OLS results)	2.730			1.872		
	(0.193)			(0.233)		
All (main 2SLS results)	3.484	18.83	45.956	2.613	68.61	34.165
	(0.406)	(0.00)	(0.066)	(0.562)	(0.00)	(0.412)
Brand & Benson	3.756	19.98	35.992	2.601	73.08	30.784
	(0.452)	(0.00)	(0.174)	(0.543)	(0.00)	(0.376)
Brands & Competition	3.482	18.70	42.234	2.723	75.10	32.099
	(0.400)	(0.00)	(0.086)	(0.605)	(0.00)	(0.412)
Benson & Competition	2.896	43.12	11.011	1.503	6.13	4.594
	(0.750)	(0.00)	(0.051)	(0.695)	(0.00)	(0.467)
Brand only	3.814	20.05	27.699	2.706	71.15	29.017
	(0.451)	(0.00)	(0.427)	(0.616)	(0.00)	(0.360)
Benson only	2.274	47.66	1.337	1.987	8.73	0.045
	(1.040)	(0.00)	(0.248)	(0.762)	(0.00)	(0.832)
Competition only	2.966	49.95	10.186	1.558	9.24	3.372
	(0.913)	(0.00)	(0.017)	(0.851)	(0.00)	(0.338)

Note: Table shows 2SLS estimates for the fuel-switching price elasticity estimated using different sets of instruments. Table also presents first-stage F-statistics for weak instruments (with robust p-values in parentheses) and Hansen's J-statistics for overidentifying restrictions (with robust p-values in parentheses). See text for details.

5.3.3 Variable elasticity estimates

In addition to the indirect tests above, I estimate a variable elasticity function directly. Because OLS does not appear to be severely biased in my application, I use OLS to estimate increasingly flexible polynomial, cubic spline, and non-parametric approximations for $F(\cdot)$. Coefficient estimates on the covariates are similar to those in table 2 above, so I focus here on the fuel-switching elasticities. For consistency with the above results, I impose $\alpha = -0.20$; recall from above that the magnitudes of the fuel-switching response and imposed α move in opposite directions.

Figures 7(a)–(b) plot the elasticity function based on a cubic polynomial approximation, estimated using the OLS fixed-effects and first-difference estimators. Elasticities decline slightly in magnitude as the ratio of gasoline to ethanol prices increases, but there is little curvature in the elasticity function. Because the cubic model would have difficulty revealing sharp peaks in the elasticity function, I also estimated more flexible polynomial approximations. Elasticities based on the higher-order polynomial approximations reveal additional non-linearities but are not statis-



(a) Fixed-effects polynomial

(b) First-differences polynomial

(c) Cubic-spline

(d) Semi-parametric

Figure 7: Estimated fuel-switching elasticities

Note: Figure shows the fuel-switching elasticity of aggregate ethanol demand based on fixed-effects polynomial, first-differences polynomial, cubic-spline, and semi-parametric estimation. Solid gray lines are 95% confidence intervals for estimates based on cubic polynomials, cubic spline with knots every 0.20, and semi-parametric estimates with bandwidth 0.06. Confidence intervals for semi-parametric estimates are based on the standard errors from local polynomial regressions. See text for details.

tically different from the more restrictive cubic estimates.

Polynomial approximations are sensitive to the number of terms and to outliers, and the function in one region is sensitive to the fit in other regions. To test the performance of the polynomials, I also estimated regression (11) using a cubic-spline approximation with knots at relative price intervals of 0.05, 0.10, and 0.20 (prior to taking logarithms), while controlling for station effects using fixed-effects estimation. Cubic-spline approximations are more flexible than polynomials and less sensitive to outliers, but they are sensitive to the number and placement of knots. Finally, I estimated the model semi-parametrically using Yatchew's (1997) estimator for the partial linear model with a bandwidth of 0.03 and 0.06.[26] Semi-parametric estimators give more detailed local approximations, but estimates are sensitive to the choice of bandwidth. See Hausman and Newey (1995) for a discussion of these tradeoffs as applied to gasoline demand.

Figures 7(c)–(d) plot the elasticity estimates from the cubic-spline and semi-parametric approaches. For neither the cubic-spline nor the semi-parametric approaches am I able to reject the least flexible of the functional forms, although the more flexible semi-parametric approach reveals additional nonlinearities that the polynomial and cubic-spline approaches obscure.

5.3.4 Preference heterogeneity

My results all point to heterogeneous preferences for ethanol. First, relative prices vary considerably across stations and over time, which would not happen if preferences were literally homogenous, since retailers would be forced to set a single price. At a minimum, fuel-switching behavior extends over a wide range of relative prices where ethanol is discounted 0%-25% below gasoline. These results suggest (1) that some consumers are willing to pay a per-mile premium for ethanol, and (2) that preferences are actually quite diffuse among households that choose ethanol at observed prices. Second, while my 2SLS elasticity estimates are large in magnitude, they are not

[26] For a general partial-linear model given by:

$$y_t = f(x_t) + Z_t \beta + \varepsilon_t,$$

Yatchew's procedure entails: (1) sorting the data by x_t, (2) differencing the data to remove the non-linear component $f(x_t)$ under the assumption that $f(x_t) \approx f(x_s)$ for $x_t \approx x_s$, (3) estimating β parametrically on the differenced data, (4) subtracting the predicted value of this parametric regression from the original dependent variable to yield $y_t - Z_t'\hat{\beta}$, and finally (5) regressing $y_t - Z_t'\hat{\beta}$ on x_t non-parametically using any number of non-parametric regression techniques. I employ tenth-order differencing using Yatchew's (1998) optimal differencing weights, which improves efficiency to within 5% of Robinson's (1988) fully efficient procedure. I control for station effects using station dummy variables. Several hundred observations share the same relative price as another observation, which means the results may be sensitive to how sorting ties are broken. I therefore estimate the first stage 50 separate times, breaking ties randomly in each trial, and take the mean coefficient values from these trials as my first-stage estimates. In practice, my estimates vary negligibly across trials. I then estimate the non-parametric portion of the model using local polynomial regression, which has attractive properties in the extremes of the data. Polynomials also yield intuitive and convenient estimates for first derivatives. I use a quadratic local polynomial, which is appropriate for estimating first derivatives (Fan and Gijbels 1996), and an Epanechnikov kernel weighting function. I calculate an "optimal" bandwidth of 0.03 using a rule-of-thumb approximation (Fan and Gijbels 1996, p.111).

nearly as large as they would be if household preferences for ethanol were more homogeneous, as in figure 1(a). Third, my linear 2SLS estimates are for the most part consistent across different sets of instruments, which is consistent with a roughly constant elasticity function. Fourth and finally, when I estimate flexible elasticity functions directly, I also find that elasticities are roughly constant. These results are all consistent with a flat elasticity function and heterogeneous preferences for ethanol, as in figure 1(b) above.

Unfortunately, because I rarely observe ethanol discounted less than 0% or more than 25% below gasoline, I am unable to estimate the elasticity function or say anything definitive about preferences in those regions. I can, however, estimate the distribution of preferences over the range of observed prices. Rough calculations suggest that about 13% of flexible-fuel owners chose ethanol during my sample period when the price ratio averaged 1.14. Thus, I can impose $H(1.14) = 0.13$ and use any one of my elasticity estimates to reveal the rest of the distribution (i.e., by integrating the elasticity function with respect to relative prices), as detailed in the appendix.

One concern is that my estimates, which reflect preferences for ethanol in a particular time and place, may not be appropriate for out-of-sample simulations. I tested whether preferences have changed over time (e.g., due to advertising campaigns late in the sample period) by estimating the model separately on data for 1997–2003 and 2004–2006. I also tested whether preferences vary cross-sectionally (e.g., with the importance of agriculture in the local economy) by estimating the model separately for stations in the Twin Cities and stations in greater Minnesota. Fixed-effects polynomial and cubic-spline estimates show no significant differences over time or across geography, however, which suggests that efforts to make the results more nationally representative would probably not alter my main findings appreciably.

One final concern is that some of the heterogeneity I observe derives not from variation in preferences per se but from variation in ethanol's relative convenience. If so, then my estimates may give inaccurate predictions when I simulate the effects of an ethanol standard, which would presumably expand ethanol's availability beyond current levels. In reality, while there are relatively few ethanol stations statewide, if a given town has an ethanol station at all, it is typically located near other gasoline stations. Thus, while some consumers may be driving out of their way for ethanol, most are probably not driving very far. Hence, correcting for variation in ethanol's convenience also would not alter my main conclusions appreciably.[27]

6 Policy simulation

I use my model and estimates to simulate the effects of an ethanol content standard, which mandates that denatured ethanol comprise a minimum fraction of the overall gasoline supply. The simulation model is necessarily stylized and intended to highlight the importance of modeling heterogeneous preferences for ethanol. I simulate 15% and 25% standards. The 25% standard is consistent with the federal RFS of 36 billion gallons annually for 2022, which represents about 25% of current gasoline consumption.[28] I assume in my simulations that industry complies with the standard by increasing the market share of retail ethanol, although blending with regular gasoline in other ratios would also be a potential compliance strategy.[29] My model and estimates could also be used to evaluate other government policies to promote retail ethanol.

I assume in my simulations that, for price ratios less than 1.35, preferences follow the cdf implied by my 2SLS fixed-effects estimates. Rather than extrapolate forward out of sample, however, I impose that the remaining mass of households share the same fuel-switching price ratio of 1.35, which is the average ratio of gasoline to ethanol mileage and therefore consistent with most con-

[27] Within a five-mile radius, 60% of competing stations are within 0.5 miles and half are within 1 mile. At a time cost of $15 per hour, travel speed of 30 miles per hour, ethanol cost of $1.75 per gallon, fuel economy of 20 miles per gallon, and refueling rate of 15 gallons, traveling one mile round-trip out of the way for ethanol (the median) would add only $0.04 to the effective price of ethanol. Adjusting for this extra distance would (given the forgoing assumptions) shift a price ratio of 1.15 to 1.12.

[28] Recall that while this standard mandates a minimum quantity of renewable fuel, the EPA rulemaking that implements the standard sets a minimum percentage of renewable fuel in each compliance period. Gasoline consumption is not projected to increase much in the coming decades.

[29] Although industry has thus far been blending ethanol with gasoline to comply, this strategy will soon be limited by the fact that regular gasoline vehicles cannot run on ethanol blends higher than 10%. Industry recently requested a waiver from EPA that would allow ethanol blends higher than 10% in regular gasoline; it is unclear whether this waiver will be approved and whether automakers would similarly modify vehicle warranties. Given that a sizeable fraction of households are willing to pay a per-mile premium for ethanol, however, it is possible that industry would eventually find it profitable to differentiate between the two fuels and recover costs by charging households with flexible-fuel vehicles a higher price for ethanol.

sumers minimizing fuel costs. In effect, I assume that the distribution of preferences has a mass point at 1.35 and a long left tail consistent with my estimates. For comparison to previous analyses that assume identical preferences, I also simulate the standards assuming that all households are massed at 1.35. I close the model by adding a supply side, drawing on previous work by Holland et al. (2008). I numerically search for retail fuel prices and a shadow value on the ethanol content constraint such that households maximize utility, suppliers maximize profits, industry complies with the constraint, and markets clear. See table 5 and the appendix for further details on the simulation.

Table 5 presents the simulation results. Scenario 1 assumes that households have nearly identical preferences based on ethanol's mileage relative to gasoline. This constrains the equilibrium price ratio under the standard to nearly equal the mileage ratio of 1.35. A 15% ethanol content standard reduces gasoline consumption by about 12% and reduces carbon dioxide emissions by about 4%. The policy is costly, however, at $12 billion annually. I calculate total costs based on changes in consumer surplus and producer surplus, with changes in tax revenue (i.e., fuel taxes net of the federal ethanol subsidy) entering as lump-sum transfers. These impacts are about twice as high for the 25% standard.

Scenario 2 assumes that households are heterogeneous as implied by my estimates. After accounting for the fact that some households prefer ethanol, the surplus cost of a 15% ethanol content standard falls by 10%. Costs are lower in scenario 2 because households with strong preferences can be induced to purchase ethanol with less severe distortion of market prices, as evidenced by the lower equilibrium price ratio of 1.1.[30] For the 25% standard, surplus costs are only 3% lower than in scenario 1. A fairly high price ratio of 1.31 is needed to comply with the standard, and so price distortion is nearly as high. The fuel supply industry benefits quite substantially under the policy, although producer surplus in the table does not distinguish between

[30] In fact, baseline ethanol consumption is actually higher in scenario 2. The expansion of baseline ethanol consumption above current levels occurs because I assume for the simulation that all households own flexible-fuel vehicles, whereas in reality this fraction is quite small. While conversion costs are low and falling over time, they are not zero, and so production of these vehicles derives primarily from CAFE incentives. Endogenizing flexible-fuel conversions by adding conversion costs to the analysis would reduce the difference between scenarios 1 and 2 and increase the cost of complying with an ethanol content standard.

Table 5: Simulation results

Scenario 1: Identical households		**Ethanol standard**	
	0%	**15%**	**25%**
ethanol price ($/gallon)	2.68	2.17	2.52
gasoline price ($/gallon)	2.62	2.90	3.38
gasoline / ethanol price	0.98	1.34	1.34
quantity pure ethanol (billion gallons)	4.95	21.61	35.99
quantity pure gasoline (billion gallons)	136.86	122.44	108.30
emissions (million mtCO_2)	1229.08	1185.28	1132.54
change consumer surplus (billion $)	0.00	-7.73	-21.00
change producer surplus (billion $)	0.00	3.07	13.65
change tax revenue (billion $)	0.00	-7.37	-14.59
total cost (billion $)	0.00	-12.03	-21.94
cost per gasoline saved ($/gallon)		0.83	0.77
cost per emissions reduced ($/mt$CO_2$)		274.69	227.24
Scenario 2: Heterogeneous households		**Ethanol standard**	
	0%	**15%**	**25%**
ethanol price ($/gallon)	3.29	2.51	2.50
gasoline price ($/gallon)	2.56	2.75	3.27
gasoline / ethanol price	0.78	1.10	1.31
quantity pure ethanol (billion gallons)	9.34	20.50	34.02
quantity pure gasoline (billion gallons)	129.38	116.35	101.93
emissions (million mtCO_2)	1185.10	1126.07	1066.61
change consumer surplus (billion $)	0.00	-3.11	-14.35
change producer surplus (billion $)	0.00	-1.11	6.96
change tax revenue (billion $)	0.00	-6.62	-13.96
total cost (billion $)	0.00	-10.84	-21.35
cost per gasoline saved ($/gallon)		0.83	0.78
cost per emissions reduced ($/mt$CO_2$)		183.70	180.22

Note: Scenario 1 assumes that fuel-switching price ratios follow a normal cdf with mean 1.35 and standard deviation 0.01. Scenario 2 assumes that fuel-switching price ratios follow the supremum of (a) a normal cdf with mean 1.35 and standard deviation 0.01 and (b) the cdf implied by my 2SLS fixed-effects elasticity estimates with $H(1.14) = 0.13$; this amounts to adding a long left tail to the distribution in scenario 1. All simulations assume: that every household owns a flexible-fuel vehicle; a price elasticity for individual ethanol-equivalent fuel demand of -0.20; price elasticities of 1.25 and 2.5 for pure gasoline and denatured ethanol supply; 8.8 kilograms of CO_2 emissions per gallon of gasoline; and that ethanol's energy-adjusted, life-cycle CO_2 emissions are 15% lower than gasoline. The aggregate ethanol-equivalent fuel demand function is calibrated to 2006 gasoline quantities and retail prices. Supply functions are calibrated to 2006 quantities and national-average wholesale spot prices; supply functions also include a constant marginal cost for distribution, marketing, and taxes net of subsidies. See the appendix for further details.

producers of ethanol and gasoline. Consumers and taxpayers split the remaining costs 60-40. Gasoline consumption and greenhouse gas emissions under the 15% and 25% standards fall by about the same amount as in scenario 1.

The ethanol content standard remains a costly policy, however, even after accounting for revealed preferences. Surplus costs in scenario 2 average about \$0.80 per gallon of gasoline saved. For comparison, a recent study by Harrington, Parry and Walls (2007) assumes \$0.12 per gallon for the external costs of petroleum dependence, though the studies they review estimate a range of \$0.08–\$0.50 per gallon.[31] Surplus costs in scenario 2 average more than \$200 per ton of carbon dioxide emissions avoided. Again, these costs exceed most estimates for climate damages from carbon dioxide emissions. A recent meta-analysis suggests that marginal damages are unlikely to exceed \$15 per ton of carbon dioxide emissions (Tol 2005), while even pessimistic recent estimates put marginal damages at only \$85 per ton (Stern 2006).

These estimates are sensitive to assumptions about ethanol's life-cycle emissions and other impacts. If land-use changes eat into ethanol's moderate climate benefits, as recent studies suggest is likely, the content standard could actually increase emissions. In addition, most life-cycle studies assume that ethanol plants use natural gas, while some new plants rely on coal, which is much dirtier. Lastly, ethanol consumes a lot of water, which may grow scarcer with climate change, while fertilizer and nutrient runoff from corn production are also damaging.

There are several other limitations to these results. First, my estimates reflect preferences of households in Minnesota that own flexible-fuel vehicles and live near ethanol pumps. These households may have stronger preferences. While price responses are not statistically different over time or across regions, I am unable to verify that the estimates are representative. Second, it is difficult to determine precisely the fraction of households that choose ethanol at observed prices, and I can only speculate about preferences outside my sample. Third, any heterogeneity that derives from variation in ethanol's convenience will likely diminish over time as the ethanol

[31] They include petroleum dependence costs in a comprehensive measure of gasoline-related externalities, which they estimate at \$2.20 per gallon. The majority of these costs depend on miles driven, however, and therefore hit ethanol even harder due to its poor mileage relative to gasoline.

market expands.

On the supply side, previous research has not estimated ethanol and gasoline supply elasticities as convincingly as one would hope, though this is an active area of research. Second, I do not consider preexisting distortions, such as agricultural price supports, nor do I consider other general equilibrium effects. Commodity prices were high in the sample period, however, and price floors were not binding. Finally, I do not consider new technologies that would facilitate cheap ethanol production from agricultural waste or other feedstocks. While the RFS actually mandates that a substantial fraction of the standard be met with such fuels, forcing these technologies prematurely could increase the cost of the standard. Addressing these various issues would have an ambiguous effect on overall costs, but the key qualitative point remains: accounting for heterogeneous preferences can (in this case) reduce simulated costs.

7 Conclusion

I develop a model that explicitly links aggregate demand for ethanol in a market to the distribution of household preferences for ethanol as a gasoline substitute. The model allows me to extract information about micro preferences from aggregate data on ethanol quantities and relative fuel prices. I need not observe gasoline quantities, in contrast to other methodologies that match predicted and observed market shares. I estimate the model using panel IV methods and data from a large number of retail fueling stations. My theoretical model implies that elasticities might vary dramatically with relative prices. Thus, I attempt to determine which part(s) of the demand function are weighted most heavily in my IV estimates using a heuristic approach, and I test whether different instruments (operating in different parts of the demand function) give different results. Taking a more direct approach, I also estimate elasticities that vary flexibly with relative fuel prices using semi-parametric estimation and other flexible methods. Future research could apply this model and these methodologies to estimate preferences for other goods with perfect substitutes. Imposing constant elasticities in such contexts may give misleading results.

I find that demand for ethanol as a gasoline substitute is sensitive to relative fuel prices, with elasticities of about 2.5–3.5. Price responses are considerably smaller and less variable, however, than they would be if preferences were identical. Fuel-switching behavior extends over a wide range of relative prices where ethanol is discounted 0%–25%, and there is no single price at which a large mass of consumers suddenly switches to ethanol. The results imply that some households are willing to pay a premium for ethanol and that preferences among these households are quite heterogeneous.

These results have important implications for policy. Accounting for households that prefer ethanol can cut the cost of an ethanol content standard substantially. While the typical household may require a large subsidy, households with stronger preferences choose ethanol with minimal price distortion, reducing costs in some cases. Similar intuition likely applies for policies to promote other "green" substitutes, such as renewable electricity, energy-efficient lighting and appliances, hybrid-electric vehicles, or organic foods. Researchers should take care to distinguish between average and marginal households when assessing the impacts of policy; assuming identical preferences for all households can yield misleading results and (in this case) overstate costs.

The ethanol content standard nevertheless remains a costly policy. Costs per gallon of gasoline saved or ton of carbon emissions avoided exceed most conventional estimates of external damages, even after revising the analysis in ethanol's favor. Moreover, to the extent that preferences for ethanol reflect pure altruism toward farmers, the environment, or national security, then the behavior I interpret as reducing costs is in fact only shifting costs, at least in part. Finally, some of the altruism may actually be misplaced. If land-use changes associated with growing feedstocks negate ethanol's climate benefits, or if ethanol's other side effects are not managed carefully, then the policy could actually damage the environment. Policies that tax or regulate carbon dioxide emissions directly tend to mitigate such side-effects.

References

Alves, Denisard C.O. and Rodrigo De Losso da Silveira Bueno Bueno, "Short-run, long-run and cross elasticities of gasoline demand in Brazil," *Energy Economics*, 2003, *25* (2), 191–199.

Anderson, Soren T. and James M. Sallee, "Using Loopholes to Reveal the Marginal Cost of Regulation: The Case of Fuel-Economy Standards," *American Economic Review*, 2010, *Forthcoming.*

Angrist, Joshua D., Kathryn Graddy, and Guido W. Imbens, "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish," *Review of Economic Studies*, 2000, *67* (23), 499–527.

Barnett, William A., Douglas Fisher, and Apostolos Serletis, "Consumer Theory and the Demand for Money," *Journal of Economic Literature*, December 1992, *30* (4), 2086–2119.

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

Dahl, Carol and Thomas Sterner, "Analyzing Gasoline Demand Elasticities: A Survey," *Energy Economics*, July 1991, *13* (3), 203–210.

Davis, Lucas W. and Lutz Kilian, "Estimating the Effect of a Gasoline Tax on Carbon Emissions," *Journal of Applied Econometrics*, 2010, *Forthcoming.*

Ellison, Sara Fisher, Iain Cockburn, Zvi Griliches, and Jerry Hausman, "Characteristics of Demand for Pharmaceutical Products: An Examination of Four Cephalosporins," *The RAND Journal of Economics*, Autumn 2006, *28* (3), 426–446.

Espey, Molly, "Explaining the Variation in Elasticity Estimates of Gasoline Demand in the United States: A Meta-Analysis," *Energy Journal*, 1996, *17* (3), 49–60.

_ , "Gasoline Demand Revisited: An International Meta-Analysis of Elasticities," *Energy Economics*, 1998, *20* (3), 273–295.

Fan, J. and I. Gijbels, *Local Polynomial Modelling and Its Applications*, Vol. 66 of *Monographs on Statistics and Applied Probability*, Boca Raton, FL: CRC Press, 1996.

Fargione, Joseph, Jason Hill, David Tilman, Stephen Polasky, and Peter Hawthorne, "Land Clearing and the Biofuel Carbon Debt," *Science*, February 2008, *319* (5867), 1235–1238.

Farrell, Alexander E., Richard J. Plevin, Brian T. Turner, Andrew D. Jones, Michael O'Hare, and Daniel M. Kammen, "Ethanol Can Contribute to Energy and Environmental Goals," *Science*, 2006, *311* (5760), 506–508.

Graham, Daniel J. and Stephen Glaister, "The Demand for Automobile Fuel: A Survey of Elasticities," *Journal of Transportation Economics and Policy*, January 2002, *36* (1), 1–26.

Greene, David L., "Motor fuel choice: An econometric analysis," *Transportation Research A*, 1989, *23A* (3), 243–253.

Griliches, Zvi and Jerry A. Hausman, "Errors in Variables in Panel Data," *Journal of Econometrics*, 1986, *31*, 93–118.

Harrington, Winston, Ian Parry, and Margaret Walls, "Automobile Externalities and Policies," *Journal of Economic Literature*, 2007, *45*, 374–400.

Harris Interactive, "Two-Thirds of Adult Vehicle Owners are Familiar with Flexible Fuel Vehicles and More than Half are Interested in Purchasing One," July 26 2006. http://www.harrisinteractive.com/NEWS/allnewsbydate.asp?NewsID=1074.

Hausman, Jerry, "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left," *Journal of Economic Perspectives*, Autumn 2001, *15* (4), 57–67.

Hausman, Jerry A. and Whitney K. Newey, "Nonparametric Estimation of Exact Consumers Surplus and Deadweight Loss," *Econometrica*, November 1995, *63* (6), 1445–1476.

Holland, Stephen P., Christopher R. Knittel, and Jonathan E. Hughes, "Greenhouse Gas Reductions Under Low Carbon Fuel Standards?," *American Economic Journal: Economic Policy*, 2008, *1* (1), 106–146.

Hughes, Jonathan E., Christopher R. Knittel, and Daniel Sperling, "Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand," *The Energy Journal*, January 2008, *29* (1).

Kennan, John, "Simultaneous Equations Bias in Disaggregated Econometric Models," *The Review of Economic Studies*, 1989, *56* (1), 151–156.

Kilian, Lutz, "The Economic Effects of Energy Price Shocks," *Journal of Economic Literature*, December 2008, *46* (4), 871–909.

Minnesota Department of Agriculture, "Economic Impact of the Ethanol Industry in Minnesota," Technical Report May 2003.

National Petroleum News, "Retail Market: Market Facts," *National Petroleum News*, July 2006, *98* (8), 96–109.

Nevo, Aviv, "Measuring Market Power in the Ready-to-eat Cereal Industry," *Econometrica*, March 2001, *69* (2), 307–342.

Perkins, Morgan and Sergio Barros, "Brazil Sugar Ethanol Update," Global Agriculture Information Network Report BR6001, U.S. Department of Agriculture, Foreign Agricultural Service February 2006.

Phillips, Owen R. and David P. Schutte, "Identifying profitable self-service markets: A test in gasoline retailing," *Applied Economics*, February 1988, *20* (2), 263–272.

Phoenix Automotive, "E85 Ethanol Flexible-Fuel Vehicle Survey," May 2006. Prepared for the American Lung Association of Minnesota.

Rask, Kevin N., "Clean Air and Renewable Fuels: The Market for Fuel Ethanol in the U.S. from 1984-1993," *Energy Economics*, June 1998, *20* (3), 325–345.

Robinson, Peter M., "Root-N-Consistent Semi-parametric Regression," *Econometrica*, 1988, *55*, 931–954.

Salvo, Alberto and Cristian Huse, "Is Arbitrage Tying the Price of Ethanol to that of Gasoline? Evidence from the Uptake of Flexible-Fuel Technology," July 2010. Working Paper.

Searchinger, Timothy, Ralph Heimlich, R. A. Houghton, Fengxia Dong, Amani Elobeid, Jacinto Fabiosa, Simla Tokgoz, Dermot Hayes, and Tun-Hsiang Yu, "Use of U.S. Croplands for Biofuels Increases Greenhouse Gases Through Emissions from Land-Use Change," *Science*, February 2008, *319* (5867), 1238–1240.

Solon, Gary, "Estimating Autocorrelations in Fixed-Effects Models," February 1984. NBER Technical Working Paper 32.

Stern, Nicholas, *Stern Review: The Economics of Climate Change*, Cambridge, MA: Cambridge University Press, 2006.

Tol, Richard S.J., "The marginal damage costs of carbon dioxide emissions: an assessment of the uncertainties," *Energy Policy*, November 2005, *33* (16), 622–634. 2064-2074.

U.S. Department of Energy, *Handbook for Handling, Storing, and Dispensing E85* April 2006.

U.S. Environmental Protection Agency, *Model Year 2000–2006 Fuel Economy Guide* 2000-2006.

Wooldridge, Jeffrey M., "Selection Corrections for Panel Data Models Under Conditional Mean Independence Assumptions," *Journal of Econometrics*, 1995, *68*, 115–132.

— , *Econometric Analysis of Cross Section and Panel Data*, MIT Press, 2002.

Yatchew, Adonis, "An Elementary Estimator of the Partial Linear Model," *Economics Letters*, 1997, *57*, 135–143.

_ , “Nonparametric Regression Techniques in Economics,” *Journal of Economic Literature*, June 1998, *36* (2), 669–721.

_ **and Joungyeo Angela No**, “Household Gasoline Demand in Canada,” *Econometrica*, November 2001, *69* (6), 1697–1709.

The Demand for Ethanol as a Gasoline Substitute

Appendix: Not for Publication

Soren T. Anderson

A Aggregate demand and household welfare

I show above that the aggregate ethanol demand is

$$Q_e(p_e, p_g) = \phi N H\left(\frac{p_g}{p_e}\right)\bar{q}(p_e). \tag{12}$$

Aggregate demand for gasoline reflects households that own flexible-fuel vehicles but choose gasoline, as well as households that do not own flexible-fuel vehicles. Gasoline demand for households that own flexible-fuel vehicles is given by

$$\phi N \int_{p_g/p_e}^{\infty} \frac{\bar{q}(p_g/r)}{r} dH(r). \tag{13}$$

By similar arguments aggregate demand for households that do not own flexible-fuel vehicles is

$$(1-\phi)N \int_{-\infty}^{\infty} \frac{\bar{q}(p_g/r)}{r} dH(r), \tag{14}$$

which is just the total number of households that do not own flexible-fuel vehicles multiplied by their average gasoline consumption. Here, as above, I rely on the assumption that flexible-fuel vehicles and $v(\cdot)$ are distributed independently of r (and of each other). Adding these two expressions gives aggregate gasoline demand:

$$G_e(p_e, p_g) = \phi N \int_{p_g/p_e}^{\infty} \frac{\bar{q}(p_g/r)}{r} dH(r) + (1-\phi)N \int_{-\infty}^{\infty} \frac{\bar{q}(p_g/r)}{r} dH(r). \tag{15}$$

Maximized utility for an individual household that chooses ethanol is

$$v(q(p_e)) + y - p_e q(p_e), \tag{16}$$

which holds whenever $r \leq p_g/p_e$, while a household that chooses gasoline has utility given by

$$v(q(p_g/r)) + y - p_g \frac{q(p_e/r)}{r}, \tag{17}$$

which holds whenever $r > p_g/p_e$. Because I assume that household utility is linearly separable in the composite good, and because I assume an interior solution with respect to consumption of this good, each household's utility function has dollar units. This allows me to compute average welfare:

$$\begin{aligned} \phi \left\{ \int_{-\infty}^{p_g/p_e} \left[v(q(p_e)) + y - p_e q(p_e) \right] dH(r) + \int_{p_g/p_e}^{\infty} \left[v(q(p_g/r)) + y - p_g \frac{q(p_e/r)}{r} \right] dH(r) \right\} \\ + (1-\phi) \int_{-\infty}^{\infty} \left[v(q(p_g/r)) + y - p_g \frac{q(p_e/r)}{r} \right] dH(r) \end{aligned} \tag{18}$$

where I have left the averaging over $v(\cdot)$ and $q(\cdot)$ implicit to simplify the notation. The top term is average welfare for households that own flexible-fuel vehicles weighted by the fraction of these households, and the bottom term is average welfare for households that do not own flexible-fuel vehicles weighted by the fraction of such households. Average welfare for households that own flexible-fuel vehicles reflects both households that choose ethanol as well as households that choose gasoline. Multiplying by the total number of households N gives aggregate welfare.

B Retail supply behavior

How will a retailer facing the demand functions developed above choose to set prices, and how will these prices respond to shifting costs? For an ethanol retailer located close to other retailers, competition will drive the retail price of ethanol down to marginal costs:

$$p_e = c_e, \tag{19}$$

where c_e is the marginal cost of ethanol. The equilibrium ratio of retail gasoline to ethanol prices is given by:

$$\rho^* = \frac{p_g}{c_e}, \tag{20}$$

where $\rho = p_g/p_e$ is the price ratio the retailer chooses, and where I assume for simplicity that the retail price of gasoline p_g is fixed exogenously by conditions in the retail gasoline market.[32] Changes in ethanol's cost relative to gasoline therefore transmit directly to relative retail prices:

$$\frac{\partial \rho^*}{\partial (p_g/c_e)} = 1. \tag{21}$$

When ethanol's cost relative to the price of gasoline increases, relative retail prices increase accordingly.

Marginal-cost pricing is not a particularly good model for understanding retail ethanol pricing behavior. Current retail ethanol markets reflect a peculiar mix of monopoly power and competition. Because relatively few stations offer retail ethanol, customer bases overlap only marginally, if at all, allowing ethanol retailers to operate largely as local monopolists. At the same time, these retailers compete directly with nearby gasoline stations in the broader fuels market, because flexible-fuel vehicle owners are able to switch seamlessly between ethanol and gasoline.

Consider first a monopolist ethanol retailer that only offers ethanol. The retailer chooses the price of ethanol to maximize profits:

$$\Pi(p_e; p_g) = Q_e(p_e; p_g)p_e - c_e Q_e(p_e; p_g), \tag{22}$$

[32] This assumption is consistent with the current fuel market, where relatively few stations offer ethanol and ethanol sales volumes are low relative to gasoline. This assumption would not be valid for a significantly expanded retail ethanol market.

where Π is retailer profit, which is a function of the retail prices of ethanol p_e and regular gasoline p_g, Q_e is the quantity of ethanol demanded as a function of retail prices, and c_e is the constant marginal cost of offering ethanol. I assume for simplicity that the retail price of gasoline p_g is fixed exogenously by conditions in the retail gasoline market.

The first-order condition of this maximization problem is given by:

$$Q_e + Q_e' p_e - c_e Q_e' \equiv 0, \tag{23}$$

where all derivatives are with respect to the retail price of ethanol and I have suppressed the arguments of functions for clarity. Collecting terms that contain Q_e', moving Q_e to the right-hand side, and then dividing by p_e and Q_e' on both sides yields:

$$\frac{p_e - c_e}{p_e} \equiv -\frac{Q_e}{p_e} \cdot \frac{1}{Q_e'}. \tag{24}$$

This is equivalent to

$$\mu_e \equiv -\frac{1}{\xi_e}, \tag{25}$$

where $\mu_e \equiv (p_e - c_e)/p_e$ is the percent retail markup of ethanol and ξ_e is the own-price elasticity of aggregate ethanol demand. This is the standard monopoly result where the retailer equates the percent retail markup to the negative reciprocal of the price elasticity of demand.

Now consider a monopolist ethanol retailer that also sells gasoline. Adding profits from gasoline sales to the maximization problem results in a modified first-order condition:

$$\mu_e + \left(\frac{Q_g'}{Q_e'} \cdot \frac{p_g}{p_e} \right) \mu_g \equiv -\frac{1}{\xi_e},$$

where Q_g' is the change in gasoline sales volume given a marginal increase in the price of ethanol, $\mu_g \equiv (p_g - c_g)/p_g$ is the percent retail markup of gasoline, and all other terms are as above. I again assume that retail gasoline prices are fixed by market competition. When a station's ethanol price has no effect on its gasoline sales, so that $Q_g' = 0$, the first-order condition reduces to the simple case above. When $Q_g' > 0$, however, the optimal price of ethanol is higher, because increasing the price of ethanol drives some consumers to gasoline at the same station. This incentive increases with Q_g'. The incentive to increase ethanol prices and drive consumers to gasoline also increases with gasoline markups μ_g.

In areas where overall competition is fierce, retailers that increase the price of ethanol are unlikely to capture many customers switching to gasoline, since these customers have many competing gasoline stations from which to choose. That is, Q_g' will be relatively small in magnitude, and pricing behavior will tend toward the simple case above. In areas where competition is weak, however, so that Q_g' is large in magnitude, the incentive to increase the price of ethanol and could be quite strong.

I now return to the case of a monopolist ethanol retailer that only sells ethanol. Restating this retailer's first-order condition in terms of the price ratio ρ by making the substitutions $p_e = p_g/\rho$ and $\xi_e = -\xi_g + \xi_f$ yields:

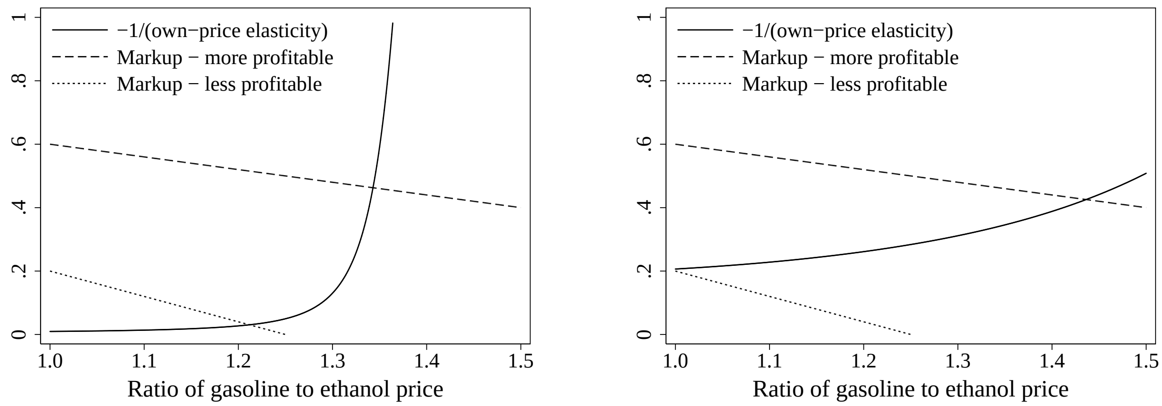
$$1 - \frac{\rho}{p_g/c_e} \equiv -\frac{1}{-\xi_g + \xi_f}, \tag{26}$$

where ρ is the price ratio the retailer chooses. Assuming that the price elasticity of individual ethanol-equivalent fuel demand ξ_f is constant, the implicit function theorem gives the following comparative static for the impact of a change in ethanol's relative cost on the profit-maximizing price ratio:

$$\frac{\partial \rho^*}{\partial (p_g/c_e)} = \frac{\rho^*}{\frac{p_g}{c_e} - \left(\frac{p_g/c_e}{-\xi_g + \xi_f}\right)^2 \xi_g'} > 0, \tag{27}$$

where ρ^* is the profit-maximizing price ratio and the inequality assumes that $\xi_g' < 0$ at the optimum. Recall that ξ_f is constant by assumption and that ξ_g and ξ_g' only depend on relative prices.

Expression (27) implies that changes in relative costs will have their largest impact on relative retail prices when the fuel-switching elasticity is roughly constant near the optimum, so that ξ_g' is close to zero. In contrast, when the elasticity is highly variable near the optimum, which indicates a large concentration of households near that same fuel-switching price ratio, ξ_g' will be large in magnitude and relative prices will be less responsive to changes in ethanol's costs. In the extreme



(a) $r \sim N(1.25, 0.05)$

(b) $r \sim N(1.25, 0.30)$

Figure 8: Profit-maximizing retail price ratio

Note: Figure illustrates profit-maximizing price ratios for a monopolist ethanol retailer. Profit-maximizing price ratios are given by intersection of markups and the negative reciprocal of the own-price elasticity, as in equation (25). More profitable and less profitable cases assume that marginal ethanol costs are 60% and 80% the retail price of gasoline. Elasticity functions assume a constant ethanol-equivalent fuel price elasticity of -0.25.

case where households have identical preferences for ethanol, ξ_g' will be infinitely large in magnitude and relative prices will be invariant to underlying costs. Retailers will be reluctant to raise ethanol prices when costs increase, lest they drive all consumers to gasoline. At the same time, retailers will have no incentive to reduce prices when costs fall, because lowering prices will not stimulate any additional demand.

Figure 8 illustrates this first-order condition and comparative static for two hypothetical fuel-switching elasticity functions, where I have set the elasticity of individual ethanol-equivalent fuel demand to a constant -0.20. The figures illustrate that when household preferences are nearly homogeneous, so that price elasticities are highly variable, as in figure 8(a), the profit-maximizing price ratio is insensitive to changes in relative costs. When household preferences are more diffuse, however, so that price elasticities are less variable, as in figure 8(b), shifts in relative costs lead to large changes in the profit-maximizing price ratio.

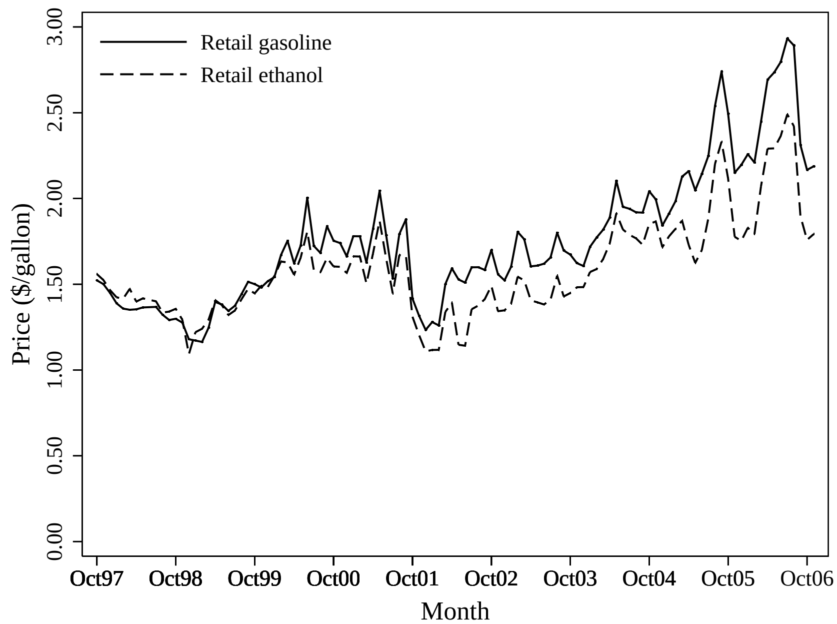


Figure 9: Retail fuel prices

Note: Retail ethanol price is the monthly volume-weighted average retail price of ethanol at reporting stations in Minnesota. Retail gasoline price is the monthly county-average retail price of regular gasoline, weighted by retail ethanol sales volumes at these same stations. Prices are in 2006 dollars.

C Aggregate price trends

Figure 9 plots average retail ethanol and regular gasoline prices from October 1997 through November 2006. Average ethanol prices track regular gasoline prices closely, albeit at a noticeable discount for most of the period.

Figure 10 plots average wholesale prices for the same time period. Wholesale spot prices for denatured ethanol do not always track wholesale gasoline prices closely. This is perhaps not surprising, given that demand for denatured ethanol derives largely from its role as a complement to gasoline production and less from its role as a gasoline substitute. Opportunities for direct substitution do exist, however, and large price differences can create strong incentives for substitution, which is one reason that wholesale ethanol prices track wholesale gasoline prices broadly over time. This is particularly evident in the fall of 2005, when ethanol helped ease gasoline shortfalls

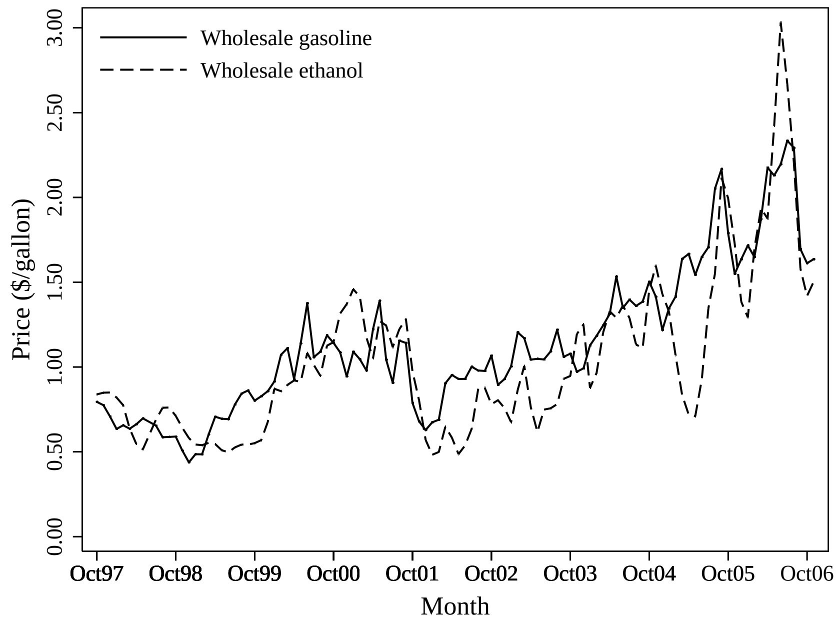


Figure 10: Wholesale fuel prices

Note: Wholesale ethanol price is a weighted average of the spot price for denatured ethanol in Minneapolis and Fargo, less the federal ethanol blending tax credit. Wholesale gasoline price is the Minnesota volume-weighted average rack price. Prices are in 2006 dollars.

after Hurricanes Katrina and Rita knocked out Gulf Coast petroleum refineries and distribution pipelines. Ethanol prices were low relative to gasoline in the first half of 2005 due to a glut of ethanol. Ethanol prices then spiked to equal gasoline prices as ethanol substituted for gasoline after the hurricanes. Ethanol's margin relative to gasoline eventually returned to pre-hurricane levels as refineries and pipelines came back on line and as imports of refined gasoline arrived from abroad.

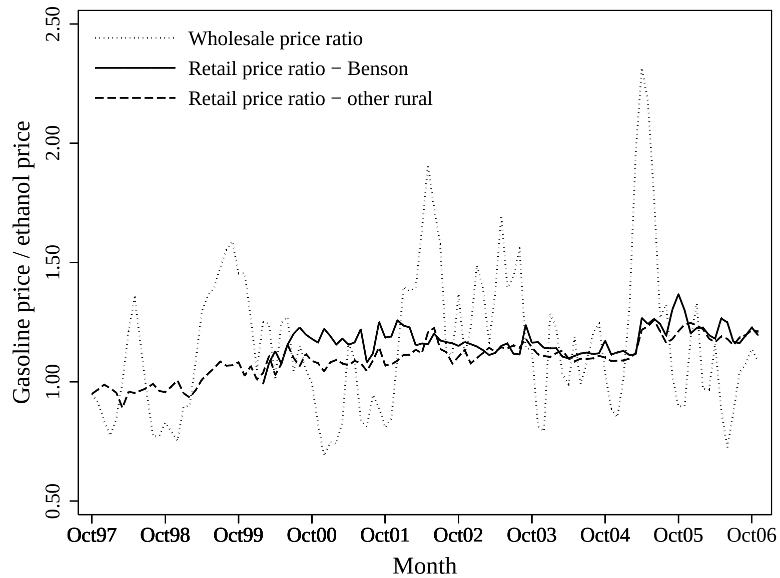
A second reason that wholesale prices track broadly is that ethanol and a petroleum-based chemical fuel additive called methyl tertiary-butyl ether (MTBE) are close substitutes in some U.S. regions during much of this time period, creating an avenue for petroleum prices to correlate indirectly with ethanol prices. The importance of this substitution is most evident in the first half of 2006, when fuel suppliers quit using MTBE due to concerns about potential groundwater contamination. Prices surged as ethanol filled the gap left by this key substitute. Ethanol prices fell in the summer months as ethanol refiners scaled up production and as fuel distributors resolved logistical difficulties in transporting ethanol from refineries in the Midwest, where ethanol is produced, to markets on the coasts, where MTBE had previously held a large market share.

Figure 5(a) above demonstrates that large fluctuations in relative wholesale prices correlate with comparatively small changes in retail prices. Note that the scale for the wholesale price ratio in figure 5(a) above is five times as large as the scale for the retail price ratio. What explains this behavior? The natural assumption is that ethanol retailers are pricing ethanol based primarily on what flexible-fuel vehicle owners are willing to pay, relative to gasoline, as opposed to what the fuel costs. As I show above in appendix section B, when the elasticity is highly variable and retailers are monopolists, the relative price of ethanol will be insensitive to changes in ethanol's costs relative to gasoline. The pricing behavior in figure 5(a) is therefore consistent with a highly variable elasticity function.

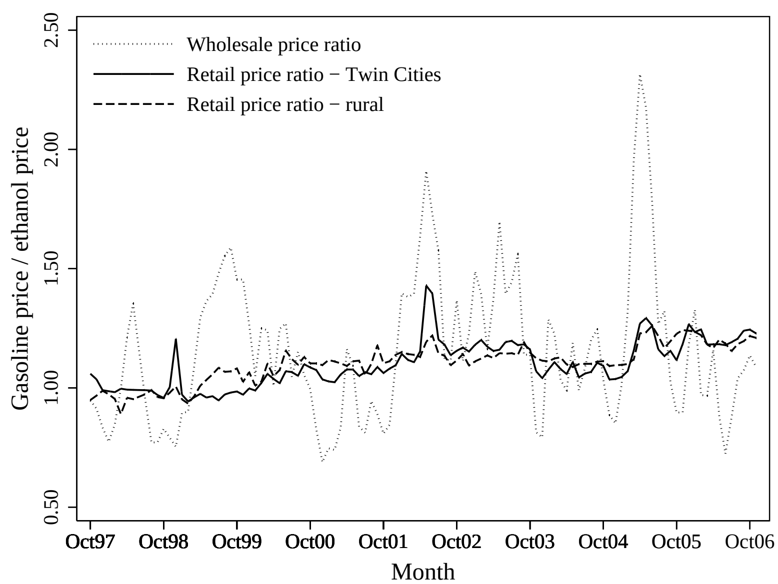
There are alternative explanations. Above, I described rule-of-thumb pricing strategies, supply relationships that mechanically tie retail ethanol prices to gasoline, and pricing formulae for long-term ethanol contracts, which could all lead retail prices to be less variable than observed spot prices. In addition, station owners may have an incentive to reduce price volatility by maintaining retail ethanol prices that are more consistent with the long-run relationship between gasoline prices and ethanol costs. Short-term profits may suffer, but this strategy helps maintain a consistent customer base. Indeed, several industry representatives I spoke with indicated that some retailers were actually pricing ethanol below costs in late 2005 and early 2006. Ethanol costs were high relative to gasoline, due to the discontinuation of MTBE, but some retailers were willing to incur temporary losses to maintain favorable relationships with their customers.

D Evidence of cross-sectional variation in pricing behavior

As I describe in the main text, about one-third of ethanol retailers in Minnesota purchase ethanol directly from an ethanol refinery in Benson, which is located in the southwestern part of the state.



(a) Benson area vs. other rural



(b) Twin Cities vs. rural

Figure 11: Relative wholesale prices and relative retail fuel prices

Note: Top figure shows relative retail prices for stations in counties within 50 miles of Benson and for other rural counties. Bottom figure shows relative retail prices for stations in Twin Cities counties and for stations in rural counties.

Throughout the entire sample period, this refinery supplied retail ethanol at a fixed nominal discount to the spot price of regular gasoline. The ethanol retailers, in turn, agreed to price retail ethanol at the same discount below regular gasoline at their stations. When retail ethanol prices are tied directly to the price of gasoline, relative prices will be less responsive to changes in ethanol's relative cost. This is apparent in figure 11(a), which plots relative retail prices for stations located in counties within 50 miles of the Benson refinery, which are most likely to have contracts with this refinery, and for stations located in other counties outside the Twin Cities. In 2000-2001, when wholesale ethanol costs were high relative to gasoline, stations near Benson priced ethanol at a larger percent discount. This happened again in late 2003-2004 and at times in late 2005-2006.

Figure 11(b) plots relative retail prices for stations located inside and outside the Twin Cities, where the density of retail ethanol stations is higher. Stations in the Twin Cities appear to be more sensitive to changes in relative costs. When wholesale ethanol costs are low relative to gasoline, stations in the Twin Cities discount ethanol more heavily than in rural areas. When wholesale ethanol costs are high relative to gasoline, retailers in the Twin Cities do not discount ethanol as generously. This pricing behavior is consistent with retailers in the Twin Cities facing greater competition and therefore being more sensitive to changes in relative costs.

E Using elasticity estimates to reveal preferences

This section shows how to retrieve the distribution of household preferences from aggregate price repsonses. Recall that equation (7) above links the distribution of household preferences to aggregate price responses:

$$\xi_g(x) = \frac{h(x)}{H(x)}x. \tag{28}$$

Dividing both sides by x and using the first-derivative rule for logarithms gives:

$$\frac{\xi_g(x)}{x} = \frac{\partial \ln H(x)}{\partial x}. \tag{29}$$

Assume that an estimate of the elasticity function is available for price ratios ranging from r_L to r_H. Then integrating both sides with respect to x through $r > r_L$ gives:

$$\begin{aligned}\int_{r_L}^{r} \frac{\xi_g(x)}{x} dx &= \int_{r_L}^{r} \frac{\partial \ln H(x)}{\partial x} dx \\ &= \ln H(r) - \ln H(r_L). \end{aligned} \tag{30}$$

Finally, taking the exponential of both sides and rearranging yields:

$$H(r) = H(r_L) \cdot \exp\left(\int_{r_L}^{r} \frac{\xi_g(x)}{x} dx\right). \tag{31}$$

Given $H(r_L)$ and an econometric estimate of $\xi_g(x)$ over the interval $[r_L, r_H]$, equation (31) yields an estimate for the cdf of household preferences on the interval.

A boundary condition is required to solve for $H(r_L)$. Suppose one knows that the fraction of households choosing ethanol at some price ratio $r^* \in [r_L, r_H]$ is $H(r^*)$. Then it is easy to solve for $H(r_L)$:

$$H(r_L) = H(r^*) \cdot \exp\left(-\int_{r_L}^{r^*} \frac{\xi_g(x)}{x} dx\right), \tag{32}$$

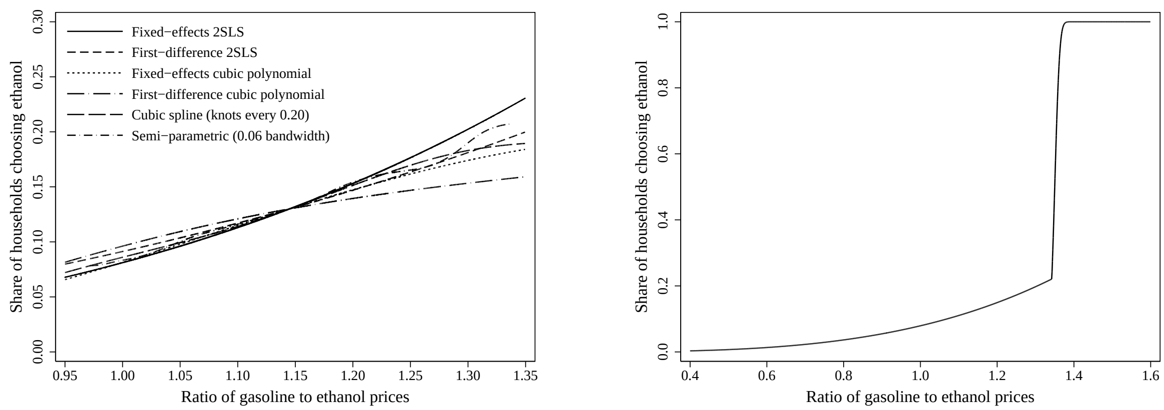
and equation (31) yields the distribution of preferences on the interval $[r_L, r_H]$.

Unfortunately, it is not possible to reveal the full distribution of household preferences for ethanol, unless one has an estimate for the elasticity function over the entire range of possible fuel-switching price ratios.

Because I rarely observe ethanol discounted less than 0% or more than 25% below gasoline, I am unable to estimate the elasticity function or say anything definitive about preferences in those regions. I can, however, estimate the distribution of preferences over the range of observed prices.

Rough calculations suggest that 5% of flexible-fuel owners in a county choose ethanol on average during the study period, when the price ratio averaged 1.14.[33] This figure likely understate

[33]There are about 3250 flexible-fuel vehicles per county on average. Assuming a flexible-fuel vehicle drives 1000 miles per month and achieves 14 miles per gallon on ethanol (the sales-weighted average based on EPA mileage estimates), this translates to $1000/14 \cdot 3250 \approx 230,000$ potential gallons per county. Actual sales average about 3350 gallons per station. With about 3.7 stations per county on average, this translates to $3350 \cdot 3.7 \approx 12,400$ gallons. Thus,



(a) Estimated cdfs

(b) Simulation cdfs

Figure 12: Estimated and simulation cdfs for fuel-switching price ratios

Note: Figure on the left shows cdfs for fuel-switching price ratios based on several alternative estimators of the fuel-switching elasticity function. Each cdf assumes that $H(1.14) = 0.13$. See text for further details. The figure on the right shows the precise cdf used in the simulations, which is based on the 2SLS fixed-effects constant-elasticity estimates for price ratios less than 1.35. See text for details.

the fraction of flexible-fuel owners that would choose ethanol if given the option, however, since not all households currently have access to the fuel, given its limited availability. Assuming, based on somewhat speculative calculations, that only 40% of the households in a county currently have access on average, then ethanol's implied market share is actually 2.5 times as high. Thus, I can impose $H(1.14) = 0.13$ and use any one my estimated elasticity functions to back out the rest of the cdf.[34] Figure 12(a) does this using the 2SLS constant-elasticity estimates in table 2 and the OLS variable-elasticity estimates from figure 7 above.

Figure 12(b) plots the precise cdf I use in my simulations below. This cdf is based on my 2SLS fixed-effects estimation of the constant-elasticity model. Following the calculations above, the cdf assumes that $H(1.14) = 0.13$. I then use my elasticity estimates to reveal the cdf for price ratios less than 1.35. Rather than extrapolate forward out of sample, however, I impose that the remaining

roughly $12,400/230,000 \approx 5\%$ of flexible-fuel owners choose ethanol.

[34]The average county in my sample is about 750 square miles. If stations drew households from within a five-mile radius, which is about 80 square miles, then 3.7 non-overlapping stations per county would cover about $3.7 \cdot 80 \approx 300$ square miles. If households were evenly distributed, then about $300/750 = 40\%$, or roughly half of consumers would have access. The fraction with access would be higher if households were more concentrated near stations and lower if the market radius was smaller than five miles.

mass of households share nearly the same fuel-switching price ratio of 1.35, which is the average ratio of gasoline to ethanol mileage for flexible-fuel vehicles and therefore consistent with most consumers minimizing fuel costs. While a large mass of households at 1.35 would imply retailers could have dramatically increased revenues by lowering the relative price of ethanol during the study period, this rarely would have been profitable given wholesale ethanol costs. Note that a price ratio of 1.35 is roughly equivalent to a 25% ethanol discount.

F Simulation details

F.1 Minimum ethanol content standard

An ethanol content standard mandates that denatured ethanol comprise a minimum fraction of the overall fuel supply:

$$\frac{\pi_e Q_e + \pi_g Q_g}{Q_e + Q_g} \geq \sigma, \tag{33}$$

where Q_e and Q_g are the aggregate retail quantities of ethanol and gasoline, π_e is the percent denatured ethanol content of retail ethanol, π_g is the percent denatured ethanol content of retail gasoline, and σ is the minimum fraction of denatured ethanol in the fuel supply as mandated by the ethanol content standard. I assume that $\pi_e \geq \sigma \geq \pi_g$, where the leftmost inequality guarantees that the ethanol content standard is technically achievable, and the rightmost inequality implies that the standard is not met trivially for any combination of fuels. Rearranging the inequality demonstrates that the standard is equivalent to a minimum market share for retail ethanol:

$$\frac{Q_e}{Q_g} \geq -\frac{\pi_g - \sigma}{\pi_e - \sigma} \tag{34}$$

An ethanol content standard is therefore identical to any fuel performance standard that implicitly mandates a minimum market share requirement for ethanol, including a low-carbon fuel standard met through increased ethanol production.

F.2 Model of the fuels market

Following Holland et al. (2008) I assume that a competitive fuel supply industry maximizes profits given by:

$$p_e Q_e + p_g Q_g - C(Q_e, Q_g) + \lambda[\pi_e Q_e + \pi_g Q_g - \sigma(Q_e + Q_g)], \tag{35}$$

where p_e and p_g are the retail prices of ethanol and regular gasoline, Q_e and Q_g are the aggregate retail quantities of ethanol and regular gasoline, $C(\cdot,\cdot)$ is the fuel industry's cost function, which is increasing in both arguments and convex, λ is the shadow value of the ethanol content constraint, and π_e and π_g are as above. Note that the total quantity of denatured ethanol required to produce the given retail quantities is $\pi_e Q_e + \pi_g Q_g$, while the total quantity of pure gasoline is $(1-\pi_e)Q_e + (1-\pi_g)Q_g$. The cost function reflects denatured ethanol and gasoline refining and distribution costs, as well as the costs of blending, distribution, and retail marketing. The cost function also reflects retail fuel taxes, as well as subsidies for denatured ethanol blending.

The first-order conditions from the fuel supplier profit maximization problem and the household utility maximization problem above together characterize market equilibrium:

$$v'(e) = \frac{\partial C(Q_e, Q_g)}{\partial Q_e} + \lambda[\sigma - \pi_e], \tag{36}$$

$$v'(rg)r = \frac{\partial C(Q_e, Q_g)}{\partial Q_g} + \lambda[\sigma - \pi_g], \tag{37}$$

and

$$\lambda[\pi_e Q_e + \pi_g Q_g - \sigma(Q_e + Q_g)] = 0, \tag{38}$$

where $\lambda \geq 0$ and where I have assumed that $v(\cdot)$ is identical for all households. The first condition holds for all consumers with $r \leq p_g/p_e$ who choose ethanol and the second condition holds for all consumers with $r > p_g/p_e$ who choose gasoline. These equilibrium conditions state that each household's marginal willingness to pay for fuel equals the fuel supply industry's marginal cost. The third condition is that either the ethanol content constraint is binding or that the shadow value of the constraint is zero.

The ethanol content standard gives an implicit subsidy of $\lambda[\pi_e - \sigma]$ for the production of retail ethanol, because the denatured ethanol content of retail ethanol exceeds the standard. Conversely, the standard imposes an implicit tax of $\lambda[\sigma - \pi_g]$ on the production of retail gasoline, because the denatured ethanol content of gasoline is less than the standard. The ultimate effect of the standard on equilibrium fuel quantities depends on the stringency of the standard, the fuel industry's cost function, the household's ethanol-equivalent fuel demand function, and the distribution of fuel-switching price ratios.

Holland et al. (2008) use a similar model to evaluate a low-carbon fuel standard met through increased ethanol production. They show that such a standard can never deliver efficient reductions in carbon dioxide emissions, because the standard implicitly subsidizes ethanol while taxing gasoline. Any first-best policy must tax all fuels that contain carbon, including ethanol, based on marginal external damages. They also show that a low-carbon fuel standard might actually increase energy consumption and carbon dioxide emissions, because the fuel supply industry could meet the standard simply by increasing ethanol production. This is similar to the well-known result that a pollution performance standard may create incentives to expand output if the rate of pollution increases less than proportionally with production. These results apply equally to my analysis of an ethanol content standard.

F.3 Calibrating demand

I assume that the fuel consumption component of individual utility is of the form:

$$v(e+rg) = k^{1/\varepsilon}\frac{\varepsilon}{\varepsilon-1}(e+rg)^{\frac{\varepsilon-1}{\varepsilon}}, \tag{39}$$

so that the household's maximization problem above yields the following expression for individual ethanol-equivalent fuel demand:

$$q(p) = kp^{-\varepsilon}, \tag{40}$$

where k is a constant, p is the ethanol-equivalent price, and $-\varepsilon$ is the constant price elasticity of ethanol-equivalent fuel demand. The assumption that individual demand has a constant price elasticity is consistent with my econometric model, which imposes a constant price elasticity of individual ethanol-equivalent fuel demand. Maximized individual utility is given by

$$\frac{\varepsilon}{\varepsilon-1}kp^{1-\varepsilon}+y. \tag{41}$$

From here, it is straightforward to calculate aggregate quantities of retail ethanol and gasoline demand, as well as aggregate household welfare, based on the general expressions available in appendix section A. Given the functional form assumption above, these expressions depend on the price elasticity of individual ethanol-equivalent fuel demand $-\varepsilon$, the scale of fuel demand Nk, the fraction of households that own flexible-fuel vehicles ϕ, and the distribution of fuel-switching price ratios $H(r)$.

I calibrate $-\varepsilon = -0.2$ based previous estimates of this parameter in the literature; this is also consistent with my main econometric estimates for fuel-switching responses, which impose this same elasticity. I then calibrate Nk based on aggregate gasoline demand and average retail gasoline prices in 2006 under the assumption that $\phi = 0$. This is consistent with current market conditions where few households own flexible-fuel vehicles and those that do have virtually no access to ethanol. I then reset $\phi = 1$ for the simulations. Simulations therefore reflect market conditions in a hypothetical world where the scale of ethanol-equivalent fuel demand is equivalent to current levels but where all households own flexible-fuel vehicles. I calibrate $H(r)$ based on my econometric estimates, as described above.

F.4 Calibrating supply

I assume that marginal costs in the fuel supply industry are given by

$$\frac{\partial C(Q_e,Q_g)}{\partial Q_e} = \pi_e K_e {B_e}^{\eta_e} + (1-\pi_e) K_g {B_g}^{\eta_g} + \psi_e + \tau_e - \pi_e \theta \tag{42}$$

and

$$\frac{\partial C(Q_e,Q_g)}{\partial Q_g} = \pi_g K_e B_e{}^{\eta_e} + (1-\pi_g) K_g B_g{}^{\eta_g} + \psi_g + \tau_g - \pi_g \theta, \tag{43}$$

where: π_e and π_g are the denatured ethanol content ratios of retail ethanol and gasoline; $B_e \equiv \pi_e Q_e + \pi_g Q_g$ and $B_g \equiv (1-\pi_e)Q_e + (1-\pi_g)Q_g$ are the quantities of pure ethanol and gasoline blend stocks required to produce the retail quantities Q_e and Q_g; the functions $K_e B_e{}^{\eta_e}$ and $K_g B_g{}^{\eta_g}$ are marginal costs of denatured ethanol and gasoline production, reflecting all costs through delivery to fuel terminals, with η_e, η_g, K_e, and K_g parameters to be calibrated; ψ_e and ψ_g are the constant marginal costs of distributing fuels to retail outlets and retail marketing, to be calibrated; τ_e and τ_g are retail fuel taxes remitted by fuel retailers to state and federal governments; and θ is the federal blending subsidy for denatured ethanol.

I assume that $\pi_e = 0.85$, because retail ethanol contains 85% denatured ethanol. I calibrate $\pi_g = 0.035$, which is the fraction of denatured ethanol in gasoline in 2006. I assume 8.8 kilograms of CO_2 emissions per gallon of gasoline and that replacing gasoline with ethanol reduces CO_2 emissions by 15% on an energy-adjusted basis. I assume that the constant price elasticity of denatured ethanol supply is $1/\eta_e = 2.5$ and that the price elasticity of gasoline supply is $1/\eta_g = 1.25$, which are the midpoints of the ranges considered by Holland et al. (2008) based on their reading of the previous literature. I then calibrate K_e and K_g based on 2006 production quantities and wholesale spot prices for denatured ethanol and unblended gasoline. I calibrate distribution and marketing costs $\psi_e = \psi_g = \$0.16$ as the differential between average wholesale prices for retail gasoline and average pre-tax retail prices. I calibrate $\tau_e = \tau_g = \$0.50$ as the average differential between pre-tax and tax-inclusive retail prices. Finally, I calibrate $\theta = \$0.51$, which is the current federal subsidy for denatured ethanol blending in 2006.

F.5 Numerical solution algorithm

The numerical solution algorithm is as follows:

(1) Choose an initial fuel price vector $p^0 = [p_e^0, p_g^0]$.

(2) Set initial shadow value of ethanol content constraint to zero: $\lambda = 0$.

(3) Compute quantities supplied based on initial price vector and first-order conditions from industry profit maximization problem. If fuel supply industry is not in compliance, increase λ, return to step (2), and iterate until industry is exactly in compliance with the ethanol content standard, yielding retail quantities supplied $S^0 = [S_e^0, S_g^0]$

(4) Compute retail quantities demanded based on initial price vector and first-order conditions from household maximization problem, yielding retail quantities demanded $D^0 = [D_e^0, D_g^0]$.

(5) If the markets clear, that is if

$$D^0 - S^0 = [D_e^0, D_g^0] - [S_e^0, S_g^0] = [0, 0],$$

then stop. Otherwise, update the price vector according to $p^1 = p^0 + \kappa[D^0 - S^0]$, where κ is a positive constant. This moves the price vector in a direction that reduces excess demand. In practice I decrease κ as the number of iterations increases in order to hone in on the market-clearing price vector. Return to step (1), and iterate.