Journal of Environmental Economics and Management 66 (2013) 251–279

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

journal homepage: www.elsevier.com/locate/jeem

Build it, but will they come? Evidence from consumer choice between gasoline and sugarcane ethanol

Alberto Salvo [a,*], Cristian Huse [b]

[a] Northwestern University, Kellogg School of Management, 2001 Sheridan Road, Evanston, IL 60208, United States
[b] Stockholm School of Economics, Box 6501, 113 83, Stockholm, Sweden

ARTICLE INFO

Article history:
Received 17 October 2012
Available online 6 May 2013

Keywords:
Gasoline
Ethanol
Biofuels
Consumer choice
Perfect substitutes
Differentiated products
Fossil fuels
Renewable fuels
Alternative fuel vehicles

ABSTRACT

How consumers might switch from gasoline and diesel to alternative energy sources is not known, since the availability of alternatives is currently very limited. To bridge this gap, we exploit exogenous variation in ethanol prices at Brazil's pumps and uncover substantial consumer heterogeneity in the choice between long-established gasoline and an alternative that is similarly available and usable: sugarcane ethanol. We observe roughly 20% of flexible-fuel motorists choosing gasoline when gasoline is priced 20% above ethanol in energy-adjusted terms ($/mile) and, similarly, 20% of motorists choosing ethanol when ethanol is priced 20% above gasoline. We use transaction-level data to explore "non-price" characteristics which differentiate the two goods in the minds of different groups of consumers. Our findings suggest—and a counterfactual illustrates—that switching away from gasoline *en masse*, should this be desired, would require considerable price discounts to boost voluntary adoption, in the US and elsewhere.

1. Introduction

Oil-based fossil fuels, namely gasoline and diesel, currently power the vast majority of the world's stock of light-duty vehicles. Yet, with the heavy dependence of oil supplies on a small number of nations, the search is on for alternative energy sources, ranging from biofuels to natural gas to electricity. Much of the alternative-energy policy debate has naturally centered on the "chicken and egg" problem—how to affordably distribute both an alternative power source and the vehicle that can run on it to consumers. However, as important to the policy debate as the supply side is the less understood demand side.[1] Do consumers essentially care only about energy prices at the pump or plug, perceiving competing fuels to be "perfect substitutes"?[2] Or do certain groups of consumers view alternative fuels as differentiated goods, even fuels that may seem fairly "similar" to gasoline, such as the liquid ethanol, forming tastes for non-price characteristics including range and perceived impacts on vehicle lifetime and the environment?

A likely reason for the paucity of research on consumer adoption is that motorists are most often held captive to a single energy source, so revealed preference studies cannot be conducted. Recently, Brazil's motor fuel markets have offered a window on how consumers might substitute an alternative fuel technology for established gasoline and diesel. A traditional sugar producer, in the late 1970s Brazil responded to the oil crisis by mandating ubiquitous supply of sugarcane ethanol across the country's fueling stations—infrastructure that survives to this day—and the introduction of ethanol-captive cars.

* Corresponding author. Fax: +1 847 467 1777.
E-mail address: a-salvo@kellogg.northwestern.edu (A. Salvo).

[1] For example, with regard to ethanol in the US, Corts [9] writes that "(w)hat the required discount to gas(oline) is remains a debated topic." (p. 7).
[2] This is a common modeling assumption, e.g., Holland et al. [14] and Salvo and Huse [23].

0095-0696/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jeem.2013.04.001

Twenty-five years later, in 2003, carmakers began speedily replacing single-fuel vehicles (either gasoline or ethanol) by dual-fuel "flexible-fuel vehicles" (FFVs), cars that can operate on any combination of gasoline and ethanol. By early 2010, Brazil's nine million-plus FFVs accounted for one-third of the light vehicle stock and, because FFVs were newer than their single-fuel counterparts, for a likely 40–50% of vehicle-miles traveled.

This paper exploits large variation in the local pump price of ethanol relative to gasoline that occurred over 2009/2010 by virtue of large fluctuation in world sugar prices. Since consumer-level data were not available, we ran our own revealed-preference survey, interviewing after observing 2160 FFV motorists make choices at the pump, thus taking advantage of the natural experiment that we saw coming in the aftermath of a poor sugarcane harvest in India (e.g., [25]).

Our first contribution is to document that there is not nearly as much consumer switching between fuels as one might expect when prices approach parity in energy-adjusted terms. Due to the locally retailed compositions and ethanol's lower energy content per unit volume, gasoline and ethanol are priced about equally in $ per kilometer (km) traveled when the per-liter price of ethanol, denoted p_e, reaches 70% of the per-liter price of gasoline p_g. That is, prices equalize when $p_e/p_g \simeq 0.7$, a threshold that the media regularly reports alongside contemporaneous fuel prices, and with which consumers are quite familiar.[3] Rejecting the "null hypothesis" of perfect substitutes, we observe (in-sample) roughly 20% of FFV motorists staying with gasoline when gasoline is priced 20% above ethanol in $/km terms, i.e., when $p_e/p_g = 0.7/1.2 \simeq 0.58$. Similarly, we observe roughly 20% of FFV motorists choosing ethanol when ethanol is priced 20% above gasoline in $/km, i.e., when $p_e/p_g = 0.7 \times 1.2 = 0.84$.

To illustrate, consider a subsample of 240 FFV motorists observed at 20 retail fueling stations across the city of Belo Horizonte in late January 2010. The average per-liter ratio p_e/p_g was 0.85 (with this ratio varying little across stations), translating to 0.241 Brazilian Real (R$) per km driven on gasoline against 0.294 R$/km on ethanol.[4] Despite the hefty 22% price premium, we observed 49 consumers, or 20.4% of the sample, choosing ethanol over gasoline. Importantly, ethanol had been more expensive than gasoline for *over six weeks*, and from our data we infer that the median consumer refueled once a week. We argue that consumers had ample opportunity to become informed about effective fuel prices, including: (i) local radio channels, with price comparisons across fuels losing only to traffic updates in newscast frequency, (ii) station servicemen, who fuel vehicles and can offer price advice upon request (and are not known for pushing one fuel over another), and (iii) even colleagues and relatives.

Our interpretation of such solid demand for the expensive fuel, whether ethanol or gasoline, is that a sizable segment of a population that is generally educated—as Brazil's urban FFV motorists are—looks beyond pump prices on choosing its energy source. To illustrate why this finding is important for energy planning, we conduct a counterfactual, predicting aggregate fuel shares in the event that the ethanol state sales tax in the northern state of Pará were to be lowered from its currently high level to that of São Paulo state, the main sugar producer and where the tax rate is lowest. We obtain only a modest effect on ethanol adoption from a substantial reduction in the relative ethanol price (noting that ethanol infrastructure is already in place). Demand for gasoline is likely to prove sticky.

One can point out that this response might grow over the long run if many years of, say: (i) local advertising by the sugar industry, with ethanol positioned to be "green," or about "local jobs," and (ii) lower ethanol prices from lower taxes, or favorable geography, helped reshape consumers' preferences, boosting "pro-environmental" and "pro-social" behaviors or generating persistent "home bias" (in the spirit of [5]). But our analysis serves to show that planners in Brazil and elsewhere should not expect tipping in fuel shares around the parity price point over the short or medium term, without potentially large investments in demand-shifters.

The paper's second contribution is an empirical investigation of why consumers might exhibit tastes over fuels' real or perceived non-price characteristics. We make several conjectures on why different consumers may perceive gasoline and ethanol to be imperfect substitutes, and show evidence of these conjectures holding up in the data. For example, we document that wealthier consumers or extensive commuters (e.g., highway travelers) are willing to pay more for gasoline, likely due to heterogeneous station stopping costs and gasoline's greater range. Older consumers and those voicing technological concerns are also more likely to choose the "universally used" fossil fuel over the more recent biofuel (even though ethanol-cum-FFV technology is certainly more established in Brazil than in other countries). On the other hand, consumers expressing concern for the environment or residing in sugarcane-growing states are more likely to choose the locally produced renewable fuel.

The closest paper to ours is Anderson [2], who examines substitution between gasoline and corn ethanol in certain midwestern US markets. Working with aggregate data, rather than transaction-level data like we do, he also identifies households who are willing to pay a premium for ethanol. However, the narrower relative price variation makes it harder to uncover a preference for gasoline (he does not observe ethanol priced favorably to gasoline, as we do), and the sparser availability of FFVs and ethanol poses selection challenges. Anderson and Sallee [3] show that consumers in many US states, while driving FFVs (supplied in part due to regulatory loopholes), cannot value flexible-fuel capacity as they are unable to find ethanol.

In contrast, ethanol has been a fact of life in retail fueling stations across Brazil since the 1980s. We believe that consumer selection among different fuel technologies in the primary car market is a lesser concern, in view of the speed and extent to

[3] Many consumers also base a price heuristic for ethanol on p_g which, due to government controls over fossil fuels, has been stable for many years. They compare p_e at the pump to a memorized $0.7p_g$.

[4] To follow the fuel economy math, notice that $0.294/0.241 = 1.22 \simeq 0.85/0.7$.

which automakers, starting in 2003, transitioned their models to the flex-fuel version alone—not giving new car buyers a choice between the flex engine and the earlier single-fuel technology.[5] While there is lower penetration of FFVs within certain imported vehicle segments (e.g., SUVs), the penetration of imports in Brazil remains low. A fair point pertains to the one-half of the market that drove older single-fuel, predominantly gasoline vehicles: less affluent motorists might, for example, incur lower time stopping costs. To the extent that this is a concern, our results should be conditioned on the one-half of Brazilian motorists who drive newer vehicles, as in any other demand study in which results are to be conditioned on existing consumers.

Different correlations that we document speak to different literatures. Casadesus-Masanell et al. [6] and Elfenbein and McManus [11] (also see references therein) examine whether consumers will actually pay more for substitute products perceived to be associated with "good causes" or linked to charity. We find strong evidence of this: some consumers do pay substantially more $ per km for ethanol and, when asked without judgment about their fuel choice, they spontaneously respond that they are motivated by the environment.[6] A related literature examines whether households are sensitive to their neighbors' energy conservation efforts (e.g., [4]). That population groups, older consumers in particular, resist switching away from an established product—their comfort zone—is consistent with the literature on technology adoption [22].[7]

The paper is structured as follows. Section 2 conjectures on why consumers might perceive fuels as differentiated products. Section 3 discusses our fieldwork's setting and design, and presents descriptive statistics including empirical demand curves. Section 4 analyzes consumer response by way of multinomial probit models. Section 5 reports evidence from follow-on phone interviews. Section 6 discusses policy implications and provides a price counterfactual. Section 7 further examines heterogeneity and Section 8 concludes.

2. Conjectured departures from perfect substitution

We briefly conjecture why substitution by consumers between gasoline and ethanol might be less than "perfect." To fix ideas, let k_{fi} denote the fuel economy of consumer i's FFV in km per liter of fuel $f \in \{g, e\}$. The expenditure per period (a year, say) by a consumer driving M_i km on fuel f is then $M_i p_{fi}/k_{fi}$, recalling that p_{fi} denotes the consumer price in $ per liter. Consumers choosing only based on fuel prices would pick fuel f with the lowest p_{fi}/k_{fi}, expressed in $/km. We label such behavior "perfect substitution."

Begin with the time cost of stopping to refuel. Suppose that the consumer incurs a linear cost of S_i $ per stop. If the tank is depleted and filled up every time with fuel f, the annualized stopping cost is $M_i S_i/(T_i k_{fi})$, where T_i is the vehicle's tank capacity in liters. Since $k_{ei}/k_{gi} \simeq 0.7$, this stopping cost makes consumers, all else equal, favor gasoline over ethanol. In particular, wealthier consumers might exhibit a higher stopping cost and thus be more likely to choose gasoline over ethanol relative to less affluent consumers (Conjecture 1, Table 1). Alternatively, suppose that some consumers exhibit convex, rather than linear, stopping costs. For example, consumers might be "loyal" to their local fueling station or have a favored fueling occasion, making them averse to fueling outside their favored station or occasion.[8] Highway commuters may be particularly prone to "range anxiety." Compared to consumers who do not drive much, convex stopping costs may make heavy commuters more likely to choose gasoline over ethanol since, all else equal, they are more likely to find themselves stopping to refuel outside their favored location or occasion, thus valuing the extra range afforded by a tank of gasoline (Conjecture 2, Table 1).

Now consider differential maintenance costs as perceived by motorists. Again to fix ideas, denote by C_{fi}^{mainten} consumer i's perceived (and additive) maintenance cost per km traveled on fuel f. We subsequently report some evidence that, compared with ethanol, more consumers associate gasoline with lower lifetime maintenance costs or improved performance than the other way round (Conjecture 3, Table 1). However, perceptions seem to vary considerably in terms of intensity and direction. Further, following the technology adoption literature [22], we conjecture that older, disproportionately traditionally minded consumers may be more inclined to stay with the long-established "technology" gasoline over the newer rival technology ethanol. This effect may be stronger to the extent that older motorists experienced technological glitches during the introduction of ethanol-dedicated cars in the 1980s. For example, the tanks of early ethanol cars were not lined with

[5] For example, conditional on buying any Volkswagen car model as of 2006, a consumer would acquire an FFV [23]. With the collapse in the price of electronics, a carmaker's cost upcharge in equipping a model with a flex engine relative to a single-fuel one is about US$ 100–200 [9,3], possibly not worth the cost of carrying different engines.

[6] The US EPA recently determined, based on lifecycle emissions analysis, that "(e)thanol from sugarcane complies with the applicable 50% reduction threshold (in greenhouse gas emissions) for advanced biofuels (compared to the 2005 gasoline baseline)" ([24]; parentheses added for clarity). In contrast, when it comes to local pollutants—i.e., urban air quality—rather than global emissions, Jacobson [16] finds that the use of ethanol is not superior to gasoline.

[7] Subsequent work should investigate the extent to which some expensive fuel choices owe to price misconception rather than tastes (e.g., some consumers not realizing how much more they are paying for a substitute which, had they been attentive, they would view as dominated), as is the focus in, for example, Miravete [19] and Clerides and Courty [7].

[8] We report some evidence of this below. One way to think about this is that some consumers may trust the quality of the fuel or service at their neighborhood station. Or a preference for stopping at certain times but not others may make them value range, as they have more control over when to stop.

Table 1
Conjectured departures from perfect substitution.

Conjecture 1: **Time Costs and Wealth** Time stopping costs may make wealthier consumers more likely to choose gasoline over ethanol relative to less affluent consumers
Conjecture 2: **Range and Vehicle Usage** Convex (time and other) stopping costs may make extensive commuters more likely to choose gasoline over ethanol relative to consumers who drive their vehicles less
Conjecture 3: **Technological Concerns** Consumers voicing technological concerns as the primary reason for their choice of fuel may be more likely to choose gasoline over ethanol relative to other consumers
Conjecture 4: **Age and Technology Adoption** Older consumers may be more likely to choose gasoline over ethanol relative to younger consumers
Conjecture 5: **Environmental Concerns** Consumers voicing environmental concerns as the primary reason for their choice of fuel may be more likely to choose ethanol over gasoline relative to other consumers
Conjecture 6: **Home Bias** Consumers who reside in sugarcane-growing states may be more likely to choose ethanol over gasoline relative to consumers who reside in ethanol-importing states

alcohol-resistant polymers so back then they were more prone to corrosion compared to gasoline cars (Conjecture 4, Table 1).

Next consider perceived effects on the natural environment. As reported below, when asked about "which motor fuel pollutes less and is better for the environment: gasoline, ethanol or is there no difference?," the vast majority of consumers respond ethanol. Some of these consumers may intrinsically care enough for the environment and decide to take matters into their own hands.[9] Denote by C_{fi}^{environ} consumer i's perceived *and* internalized cost to the environment from driving 1 km on fuel f. We conjecture that consumers who are more susceptible to environmental concerns are more likely to choose the renewable fuel ethanol over the fossil fuel gasoline (Conjecture 5, Table 1).

Finally, consider perceived differential effects on the local economic environment. In a similar vein to the natural environment, consumers when surveyed overwhelmingly opine that consumption of ethanol over gasoline is "better" for their home region, justifying this view, as we describe below, by arguing that ethanol, for example, "is a local product" or that its adoption "creates local jobs." To the extent that some consumers perceive greater benefits to the home economy from consuming ethanol *and* they display home bias in their fuel choices—thus internalizing "per-km benefits" $B_{ei}^{\text{home}} > B_{gi}^{\text{home}}$—they may, all else equal, favor "locally sourced" ethanol over "imported" gasoline.[10] In particular, we conjecture that, controlling for relative prices, residents of sugarcane-growing, ethanol-producing states of Brazil are more likely to choose ethanol over gasoline relative to residents in locations that import ethanol from out of state, for whom sugarcane features less prominently in the local economic landscape (Conjecture 6, Table 1).

To sum up, the annualized cost for consumer i from choosing fuel f (in the linear stopping cost case) is given by:

$$M_i\left(\frac{p_{fi}}{k_{fi}} + \frac{S_i}{T_i k_{fi}} + C_{fi}^{\text{mainten}} + C_{fi}^{\text{environ}} - B_{fi}^{\text{home}}\right) \tag{1}$$

While not meant to provide a theory of consumer choice between alternative energy sources, this expression illustrates some "taste-based" tradeoffs that (price-attentive) consumers may make on choosing between gasoline and ethanol. It is also plausible that consumers who make extensive use of their vehicles, or who are less affluent, may place more value on the price characteristic relative to the non-price attributes.

3. Exploiting a natural experiment

Opportunity. For years, we had been monitoring world sugar and local fuel markets, in the hope of exploiting variation in relative prices that is arguably exogenous to local demand shocks. Figs. 1 and 2 depict our empirical strategy.

Fig. 1 reports temporal variation over the period 2000 to 2010 in: (i) world prices for oil and sugar, in the upper panel, and (ii) gasoline and ethanol prices at the pump in the city of São Paulo, in the lower panel. (All prices are in constant Brazilian reais, R$; divide by two for rough prices in US dollars.) Two points, applicable generally across the country, should be noted. First, local fossil fuel prices are controlled by the central government, and the 2003–2010 administration did not allow the price of gasoline to fluctuate with the price of crude. To see this, notice that rising world oil prices, peaking in mid

[9] For example, a source cited in the Economist [10] "says green-minded drivers are prepared to pay a premium of about 30% over the cost of petroleum-based diesel to fill their cars with biodiesel." In our setting, driving off with ethanol hidden away in the tank is not an act of conspicuous consumption.

[10] This follows decades of pro-ethanol advertising by the sugar industry and the government. Given Brazil's recent deepwater oil discoveries, such views may change over the coming decades. Says US apple grower Mark Barrett, in the context of increased imports from China: "I believe if we had country-of-origin labeling that the consumers would buy US all the time" [20].

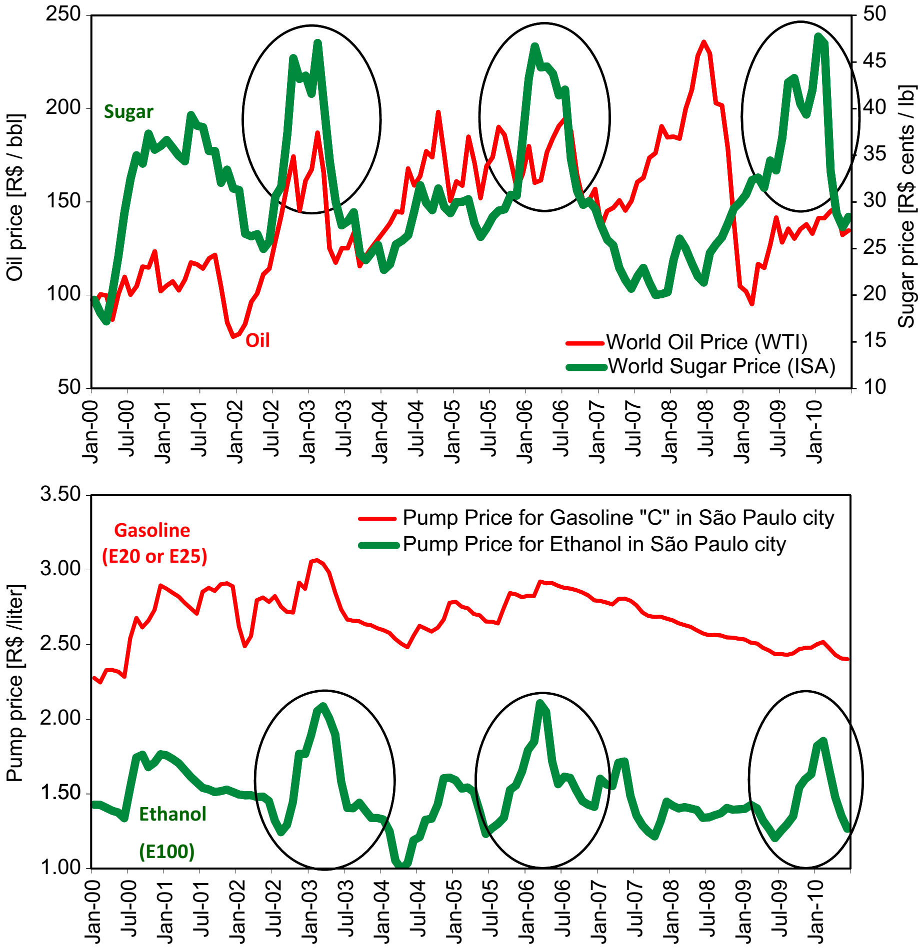

Fig. 1. The pump price of ethanol peaks when the world price of sugar peaks. Upper panel: World sugar price (ISA, R$ cents per pound) and World oil price (WTI, R$ per barrel). Lower panel: Ethanol (E100) and Gasoline (regular, E20-25) prices at the pump in the city of São Paulo (R$ per liter). All 2000–2010 prices are in constant Brazilian reais, R$, Brazil CPI base March 2010.
Sources: IBGE, EIA, ISO, BCB.

2008, were not passed through to the gas pump. Second, in contrast to fossil fuels, market forces have been at work in Brazil's sugar/ethanol industry.[11] With the opportunity cost of selling ethanol on domestic markets being given by the export price of sugar, ethanol prices at the pump have peaked every time over the past decade that world sugar prices have risen beyond a certain threshold. This is the source of relative price variation that our study exploits.

Seeing the world sugar price rally during 2009, we designed a consumer-level study and put resources on standby, ready to be deployed into the field at informative price points in time—namely, to observe and interview FFV motorists making choices at fueling stations. Instead of estimating demand from market-level data, we favored a consumer-level approach since our aim was to examine heterogeneity among subsets of the population. Further, while good-quality fuel price data exist, market-level quantity data are unavailable.[12]

Fig. 2 hones in on the 2009/2010 relative fuel price fluctuation in six major cities where we chose to observe consumers. The curves in each panel indicate percentiles of the within-city distribution of the ethanol-to-gasoline price ratio, p_e/p_g, according to large weekly samples of fueling stations from an external source, the National Agency for Oil, Biofuels and Natural Gas (ANP). Due to the ubiquitous supply of both fuels, we calculate price ratios for regular grades of ethanol and gasoline sold at the same station (and most often available at the same pump, such that drivers need not maneuver around the station to find their fuel of choice). The nine vertical lines mark the nine "city week" pairs in which we deployed representatives of a market research firm we hired. We chose cities such that three of them were capitals of sugarcane-growing, ethanol-producing states (the cities of São Paulo, Curitiba and Recife), whereas three other cities were the capitals of ethanol-importing states (Rio de Janeiro, Belo Horizonte and Porto Alegre).

[11] See Salvo and Huse [23] for a model of this industry, whose supply chain was deregulated in the 1990s. By contrast, Brazil's government controls wholesale fossil fuel prices by way of the state-controlled oil company Petrobras, a vertical monopolist all the way from exploration to refining.

[12] In particular, FFV fleet size and usage relative to older single-fuel cars, by state or city.

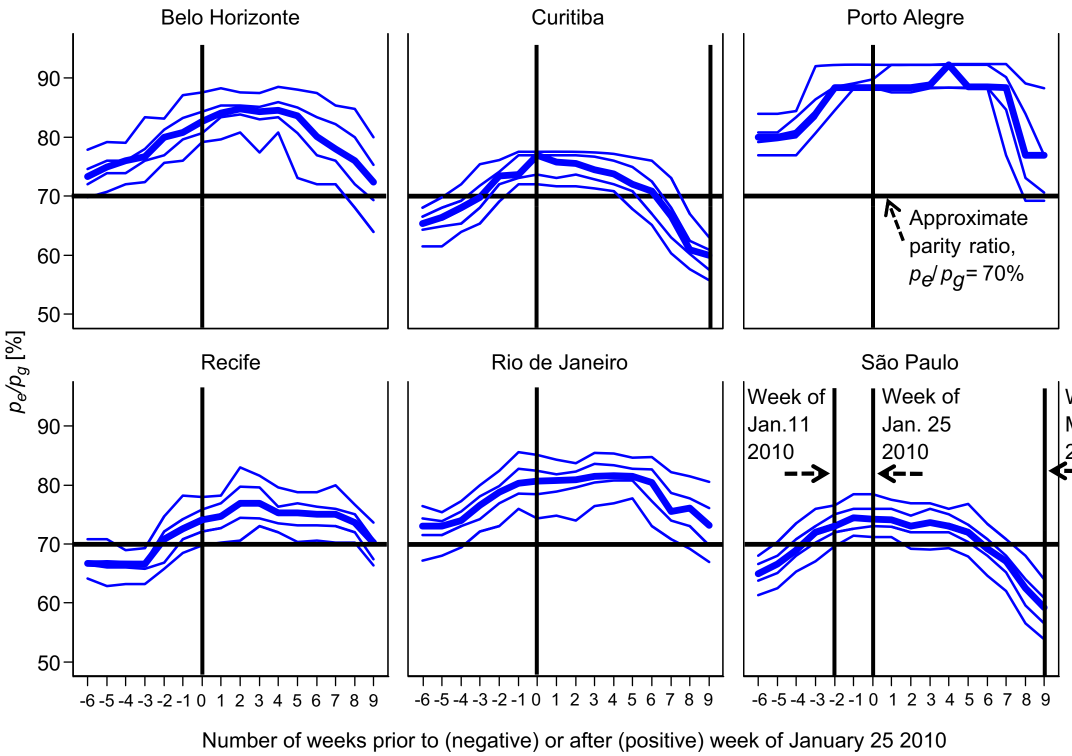

Fig. 2. Opportune price variation in the first quarter of 2010. Percentiles of the distribution across branded stations of the ethanol-to-regular-gasoline per-liter price ratio, in six cities in each of several weeks running up to and following the week of January 25, 2010. The 5th, 25th, 50th (thick), 75th and 95th percentiles of the price ratio are shown. Source: ANP's retail price database. City-weeks in our survey are marked with vertical lines.

While the p_e/p_g ratio followed the same temporal pattern in each city, rising over the weeks leading up to January 25, 2010 and falling thereafter, this price ratio at stations in ethanol-importing locations tended to exceed that in ethanol-producing ones. For example, the median p_e/p_g peaked at about 90% in Porto Alegre and 75% in São Paulo. While the wider state of São Paulo is the country's largest sugarcane grower, Porto Alegre is 2000 km from the nearest sugarcane plantation.[13] Notice also that within-city relative price dispersion was low compared to relative price variation across cities.

Perfect substitution and the media. What does this variation mean in terms of energy-adjusted prices? To illustrate, consider the fuel efficiency of a best-selling car, as measured in the laboratory, according to the National Institute for Metrology (Inmetro): a Fiat Palio ELX 1.0 2010 “Flex” (no other engines were offered) operated under an “urban driving cycle” produces 9.9 km/l of gasoline and 6.9 km/l of ethanol. For this vehicle and the local composition of fuels, $k_{ei}/k_{gi} = 6.9/9.9 \simeq 0.70$.

Indeed, this 70% relative “price parity” ratio is regularly reported in the media, particularly at times of shifting prices, not least because ANP prices for many cities are so readily accessible online. Intense media coverage of effective fuel prices includes the local radio, which Brazil's urban motorists, often stuck in traffic, spend hours listening to, becoming informed not only about traffic conditions but also receiving updates on where p_e/p_g in their local market stands relative to 70%.

A hypothetical Fiat Palio user who chooses fuel on the basis of energy-adjusted prices alone would thus have purchased gasoline in the January 2010 city-weeks of our sample—since $p_e/p_g > k_{ei}/k_{gi}$ or, equivalently, $p_g/k_{gi} < p_e/k_{ei}$. In March 2010, in São Paulo or Curitiba, this consumer would have chosen ethanol. Even in high-cost Porto Alegre, ethanol had been favorably priced relative to gasoline as recently as late 2009. As we explain below, the median k_{ei}/k_{gi} across FFVs in our field sample is 0.69, and there is little variation around this ratio.[14]

Survey design. As stated, our sample consists of nine city-weeks, including multiple city-weeks in São Paulo (three) and Curitiba (two). The first two city-weeks in São Paulo (weeks of January 11 and of January 25) exhibited similar prices but

[13] Salvo and Huse [23] show that each of the three ethanol-producing states was characterized by a share of the national sugarcane harvest that exceeded its share of national GDP (noting that ethanol mills locate close to plantations). For example, in 2005–2007 São Paulo accounted for 61% of the country's harvested sugarcane and 34% of its GDP. By contrast, the other three surveyed states were net consumers (importers) of ethanol. Thanks to a “uniform pricing” policy, p_g varied considerably less across the country, such that cross-sectional variation in p_e/p_g arose primarily from variation in p_e. In addition to differential distribution costs, some producer states supported their local sugar industry by way of a lower state sales tax on ethanol—see the counterfactual exercise.

[14] Gasoline retailed in Brazil contains a 20–25% ethanol component, respectively E20 and E25, changing occasionally by federal mandate. Inmetro publishes k_{gi} for a E22 blend, which we linearly adjust to E25 (sold in January 2010) or E20 (March 2010). Retailed ethanol is unblended E100.

were a fortnight apart: the rise in p_e/p_g petered out at 75% over January, prior to dropping. Our aim was to use these observations in an attempt to gauge short-run consumer “inertia” or “inattention,” which we conjecture to be an additional reason—on top of the “taste-based” ones enumerated in Table 1—why fuel choices might depart from perfect substitution. In the two sampled city-weeks at the end of March, in São Paulo and in Curitiba, p_e/p_g had dropped to just shy of 60%. We thus also collected data at this lower price point.

We designed each city-week subsample to consist of 20 visits to different retail fueling stations, and each station visit consisted of observing choices made by 12 FFV motorists. The sample of nine city-weeks thus amounts to $9\times20\times12=2160$ consumer-level observations, with each city-week totaling 240 observations. As we detail in Appendix A, we based our list of candidate stations on ANP's representative sample, and requested that the market research firm sample at most one station per neighborhood.

By design, we required that regular gasoline g (*gasolina comum*) and ethanol e (*álcool comum*) be available during each station visit, a constraint that in practice tended to be non-binding. Most stations also retailed “midgrade” gasoline (*gasolina aditivada*), denoted $\overline{g}$, containing “cleaning additives” but the same octane rating as the regular variety. There was no midgrade equivalent for ethanol. Very few stations sold the higher-octane “premium” gasoline variety (*gasolina premium*), which we denote by $\breve{g}$.

We instructed the field representative conducting a visit to discreetly observe each FFV motorist place his order with the station serviceman. Only once the serviceman had begun fueling, would the representative approach the consumer to collect information about him and his vehicle (consumers were prone to taking questions as they had to wait while the vehicle was being fueled). The short interview: confirmed that the car was an FFV driven on private business, which in practice is easy to spot; asked about “the main reason” for the consumer's choice of fuel on that occasion, without showing a menu of options as this might frame the response; and inquired about the consumer's age, education, and vehicle usage; among other questions.

On completing an observation, the representative would move to the next FFV motorist, by order of arrival at the station, typically having to wait for the next FFV to pull up. Station visits lasted 2.5 h on average, to complete 12 observations of short duration each, consistent with anecdotal evidence that fueling stations rarely experience customer queues.

Descriptive statistics. We keep this subsection brief and refer the reader to Appendix A for details. Fig. 3 indicates that the visited fueling stations were scattered across the six cities. Both regular gasoline and ethanol were widely available at each station—on average five nozzles dispensed g and four nozzles dispensed e. Stations did not visibly specialize in one fuel or another: larger stations tended to be equipped both with more g-nozzles *and* more e-nozzles. Importantly, “shelf space” for

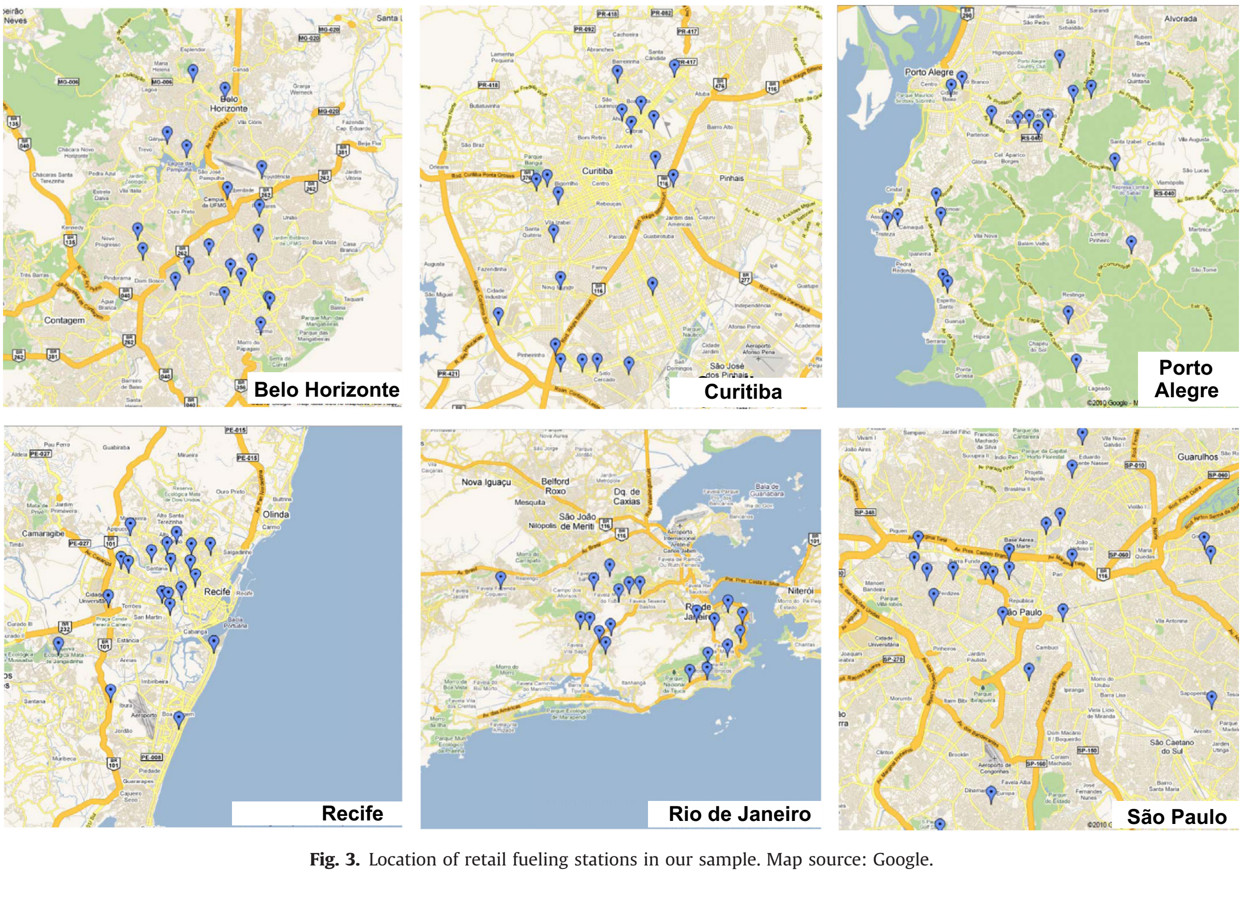

Fig. 3. Location of retail fueling stations in our sample. Map source: Google.

g versus e hardly varied between January and March visits. Midgrade gasoline $\overline{g}$ was available on 91% of station visits, retailing at a markup over g averaging 4% (and a minimum markup of 0%). With $\overline{g}$ and g having the same octane rating, we will model a consumer's choice set as $\{g, e, \overline{g}\}$ when midgrade gasoline is available at the station and $\{g, e\}$ otherwise. In the 11% of station visits for which higher-octane premium gasoline $\breve{g}$ was available, we observed only 3 consumers purchase this fuel and thus do not include it in consumers' choice sets.

Most motorists in our sample were observed—or stated—to be male (66%), middle-aged (46% were aged between 25 and 40 years, 40% between 40 and 65 years), schooled (50% stated having a college degree), and primarily motivated by “price” when purchasing fuel (68%). The median stopping frequency was once a week.[15] The majority of the sample stated that they had stopped to purchase fuel at the same station (at least) three times in a row, suggesting that consumers were not substantially “shopping around” in these markets in equilibrium.

The interviews also suggest that our station visits were *not* taking place right at the moment when consumers were beginning to switch between fuels, in which case a large mass of less-attentive consumers might still have been unaware of very recent price changes. In the subset of FFV motorists whom we observed purchasing regular gasoline, 83% stated having also purchased gasoline on both of their two immediately preceding station stops (with the caveat that this statistic is stated). Similarly, among consumers whom we observed purchasing ethanol, 78% stated having chosen ethanol on both of their two preceding stops. As shown in Fig. 2, by the week of January 25 the ethanol price hike had mostly already occurred. Similarly, by the week of March 29, p_e/p_g had already dipped below the 70% price parity threshold two to three weeks earlier.[16]

Empirical demand. Fig. 4 summarizes the choices made by FFV motorists by aggregating these to the station visit level. Markers denoting January visits are hollow and those denoting March visits are solid; further, visits to stations in São Paulo and Curitiba (multiple city-weeks) are marked with circles—hollow or solid—rather than squares. We plot p_e/p_g, the station's per-liter ethanol price relative to regular gasoline, against ethanol's overall share in the 12 choices we observed in that visit. We calculated market share for ethanol in two ways: (i) as the fraction of the 12 consumers whose fuel contained a majority ethanol by usable energy content (“vehicle km purchased”), shown in the left panel of the figure, and (ii) as the energy-weighted share of all fuel purchased by these 12 consumers that was ethanol, in the right panel.[17]

Both the unweighted and the weighted ethanol shares indicate that there is considerable consumer heterogeneity, which is the main message of our paper. Departure from perfect substitution is significant. Even among different consumers purchasing fuel at the same station on the same day, fuel choices vary. A close look at either panel reveals that facing g priced roughly 20% above e per km traveled—i.e., p_e/p_g of around $0.7/1.2 \simeq 58\%$—about one-fifth of consumers bought gasoline. Similarly, facing e priced 20% above g—i.e., p_e/p_g of $0.7 \times 1.2 = 84\%$—about one-fifth of consumers chose ethanol.

One may ask how much of this consumer heterogeneity is explained by variation in the fuel economy ratio k_{ei}/k_{gi} across FFVs, noting that the interdecile range in the surveyed sample is 0.038 around a median prediction of 0.690.[18] By controlling for “parity” differences across vehicle models, Fig. 5 shows that the answer is “not much.” To plot the figure, we compute the difference $p_{ei}/p_{gi} - k_{ei}/k_{gi}$ for each of the 2160 observations; we then collect observations in one percentage point bins and calculate the fraction of consumers who chose ethanol rather than gasoline as their dominant energy source. For example, a consumer driving a VW Gol 1.0 Flex (another popular vehicle) who we observed at a Belo Horizonte station in January faced relative prices $p_{ei}/p_{gi} \simeq 88\%$ at the pump. With a predicted $k_{ei}/k_{gi} \simeq 0.70$ for his vehicle, this consumer's choice would enter the $0.88-0.70 \simeq 18\%$ bin. Despite facing p_{fi}/k_{fi} of 0.28 R$/km on ethanol against a substantially cheaper 0.22 R$/km on regular gasoline, Fig. 5 indicates that the empirical probability that this motorist would have chosen ethanol is still no less than a sizable 10–15%.[19] This 27% ethanol price premium represents 12 R$/week, equivalent to an annualized R$ 624 expenditure difference, computed at the median vehicle usage.

Importantly, the heterogeneous consumer response depicted in Figs. 4 and 5 cannot be explained by differences in vehicle condition or average route speed. While these unobserved characteristics impact absolute fuel economy k_g and k_e in

[15] We calculate the stopping frequency from the observed quantity purchased, the stated vehicle usage and the vehicle's fuel economy under urban driving.

[16] We note that among those consumers observed purchasing gasoline in São Paulo on the week of January 11, just as p_e/p_g hit a peak, as many as 69% stated having purchased gasoline on their two preceding stops. A fortnight later, still facing peak prices, a somewhat higher 85% of gasoline consumers observed in São Paulo stated having purchased gasoline on their two preceding stops.

[17] Specifically, we compute the first, discrete-valued share as

$$s_{el}^{unweighted} = \frac{1}{12} \sum_{i \in \mathcal{O}_l} \mathbf{1}\left[q_{ei}(k_{ei}/k_{gi}) > \sum_{f \in \{g, \overline{g}, \breve{g}\}} q_{fi} \right]$$

where $\mathcal{O}_l$ is the set of consumers observed during station visit l, $\mathbf{1}[x]$ is an indicator function equal to 1 if condition x holds and 0 otherwise, and the consumer's fuel purchase in liters, denoted by q_{fi}, is adjusted for ethanol's lower km/l relative to gasoline. We do not employ the simpler condition $q_{ei} > 0$ since 2.5% of consumers in our sample purchased a “combo” of fuels on the same occasion, say R$ 30 of g and R$ 20 of e, requiring that the serviceman handle two nozzles. The second share is

$$s_{el}^{weighted} = \left(\sum_{i \in \mathcal{O}_l} q_{ei}(k_{ei}/k_{gi}) \right) \Bigg/ \sum_{i \in \mathcal{O}_l} \left(q_{ei}(k_{ei}/k_{gi}) + \sum_{f \in \{g, \overline{g}, \breve{g}\}} q_{fi} \right)$$

[18] We consider the predicted fuel economy ratio under urban driving (see Appendix A). The median prediction standard error is 0.005, i.e., the prediction is fairly tight.

[19] Had we instead plotted $p_{ei}/k_{ei} - p_{gi}/k_{gi}$ on the vertical axis of Fig. 5, the plot would look very similar. In this case, the Belo Horizonte motorist's choice would enter (say) the 0.06 R$/km bin.

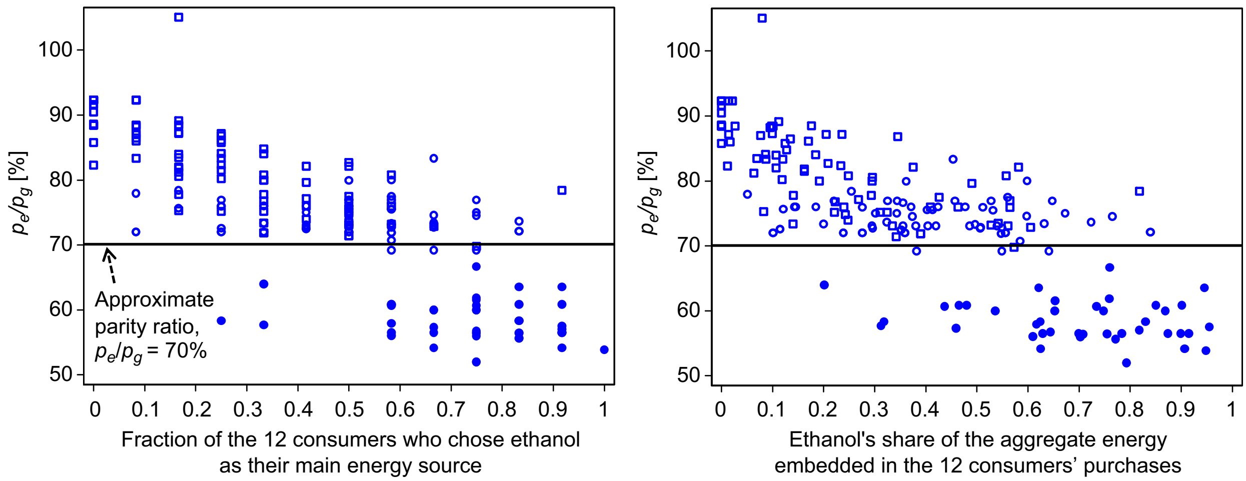

Fig. 4. Empirical demand by FFV motorists at the station level. Per-liter ethanol price relative to regular gasoline (p_e/p_g) plotted against ethanol's overall share in the 12 choices observed in each station visit. The left and right panels respectively plot the "unweighted" and "weighted"ethanol shares, as defined in the text. An observation is a station visit. Markers denoting January visits are hollow and those denoting March visits are solid; visits in São Paulo and Curitiba are marked with circles and those in other cities with squares. Source: Own survey.

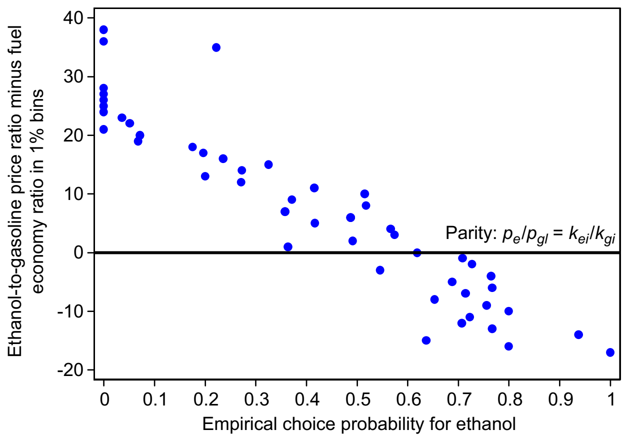

Fig. 5. Empirical choice probability for ethanol controlling for energy-adjusted relative prices. An observation is an integer category defined over 2160 choices. The vertical axis shows 1 percentage point "bins" for the difference between the ethanol-to-regular-gasoline per-liter price ratio the consumer faced at the station and the predicted ethanol-to-gasoline fuel economy ratio for his FFV. The horizontal axis reports the proportion of consumers in that bin who chose ethanol as their dominant energy source (see text). Source: Own survey.

km/l (see Appendix A), they are unlikely to materially affect the relative fuel economy k_e/k_g as this depends primarily on the relative usable energy content of the two fuels.

4. Modeling consumer choice

Given the structure of the data and the institutions, we model motorist i, observed during station visit l, as choosing fuel f to maximize

$$u_{fi} = \alpha(p_{fl}/k_{fi}) + x'_l\beta_{1f} + x'_i\beta_{2f} + \varepsilon_{fi}, \tag{2}$$

where p/k are station-vehicle specific fuel prices in R\$ per km traveled faced at the pump, vectors x_l (or x_{fl}) and x_i respectively contain other observed station and consumer/vehicle characteristics that shift choice probabilities, and unobserved idiosyncratic tastes ε follow a multivariate Normal distribution with mean zero and covariance matrix Ω, i.e., $\varepsilon \sim MVN(0, \Omega)$.[20] As mentioned, we take the consumer choice set to be $\{g, e, \overline{g}\}$ for the vast majority of stations where

[20] Thanks to their flexible properties (e.g., they do not impose "independence of irrelevant alternatives"), multinomial probit models have a long tradition in applied microeconomics, including transportation and industrial organization, e.g., [12] examine competition between cable and satellite TV using a dataset that is an order of magnitude larger than ours.

midgrade gasoline was available, and $\{g, e\}$ for the remaining stations.[21,22] Energy-adjusted prices consider predicted fuel economy under urban driving but estimates are robust to alternative assumptions. Since we do not observe the same consumer over time, we model the choice across fuels, not the "intensive margin," but we rely on the very low price elasticity of demand for private vehicle use reported in the literature (e.g., [15]). We also do not model substitution across stations, relying on the moderate within-route dispersion in *relative* prices and consumers' professed "station loyalty," as noted. In the unlikely event that on the day of our visit a station were to be cutting the price on ethanol but *not* on gasoline—raising ethanol sales by attracting passers-by who do not regularly fuel there (on top of marginal gasoline buyers from its own customer pool)—this would bias our results in the direction of *more* consumer switching between gasoline and ethanol. As we show next, we find "too little," not too much, switching between fuels compared to what one might expect on the basis of pump prices alone.[23]

Table 2 reports marginal effects, at mean values in the sample, for different baseline specifications. Due to space constraints, the table focuses on the probability of adopting ethanol over gasoline varieties. Standard errors, in parentheses, are clustered at the station-visit level. Similarly across the three specifications, we find that gender and (stated) schooling do not play a significant role in driving fuel adoption,[24] whereas increasing age is associated with the choice of gasoline. The latter correlation is captured by Conjecture 4 (Age and Technology Adoption) in Table 1. In particular, all else equal, consumers aged 65 years+ are 27 percentage points (ppt, per specification I) less likely to choose ethanol over gasoline compared to consumers aged 25 years–, who appear more comfortable with the alternative fuel. Extensive vehicle users—defined here as drivers whose stated vehicle usage places them in the upper quartile of the empirical usage distribution—also display an 8ppt lower propensity to choose ethanol, a finding that we take to be consistent with Conjecture 2 (Range and Vehicle Usage). Drivers of expensive vehicles—defined as those in the upper quartile of the survey's vehicle price distribution—are less prone to choosing the renewable fuel. Since vehicle value should reasonably proxy for income, we interpret this result as being consistent with Conjecture 1 (Time Costs and Wealth). A plausible complementary interpretation operates through heterogeneous concern for the environment (Conjecture 5, Environmental Concerns) and selection over vehicles: since expensive (larger, powerful) cars tend to burn more fuel and pollute more than cheaper vehicles,[25] "green-minded" types may favor smaller cars and locally grown sugarcane ethanol.[26] We attempt to control for the average income of a station's customers by including the aggregate value of the 12 vehicles sampled in each station visit, but this is not significant.

What does vary across the three specifications of Table 2 is the inclusion of: (i) city fixed effects, in specifications I and II, relative to III; and (ii) dummy variables indicating the main reason the consumer invokes as the basis for his fuel choice, in specification II, relative to I and III.

In the presence of city fixed effects (specifications I and II), price effects are estimated off of time variation (i.e., different weeks in January and March for São Paulo and Curitiba), as well as the moderate within-city-week dispersion in relative prices across stations (in all cities). With city-level shocks being soaked up by city fixed effects, and our proxy for the mean income of a station's customer pool controlling for other neighborhood-level shocks that may potentially influence a station's prices, the identifying assumption is that *relative* fuel prices are uncorrelated with unobserved local taste shocks that remain in the error. That is, relative prices $p_{fl}/k_{fi} - p_{f'l}/k_{f'i}$, whose within-city variation in the sample is explained primarily by fluctuation in the world sugar price, are assumed to be exogenous to $\varepsilon_f - \varepsilon_{f'}$ given the controls. We have no reason to believe that weather, income or other shocks may have shifted the demand for ethanol relative to gasoline between January and March 2010, two months that were similarly warm and close together. Another premise is that the way consumers respond to the 2009/2010 variation in ethanol prices is indicative of their response more generally and not specific to the episode. A hypothetical situation in which this would not hold is one where many consumers were aware and resentful of the impact of foreign weather shocks on local prices, having "punished" ethanol producers by switching out of ethanol in "unusually" large numbers by the time we observed them in January, and resisting the return to ethanol when we observed them again in March. This hypothetical would tend to shift Fig. 4's inverse demand for ethanol inward and make it

[21] We obtain very similar results when modeling all consumers' choice sets to be $\{g, e, \overline{g}\}$ and assuming, for the 16 stations where $\overline{g}$ was unavailable, a hypothetical price $p_{\overline{g}}$ equal to p_g at the station multiplied by the market's mean midgrade markup over regular gasoline.

[22] We place no additional restrictions on Ω beyond those necessary to identify the model. To illustrate, with $F=3$ alternatives, Ω has $F(F-1)/2-1=2$ free terms, say error variance parameter $\sigma^2_{\overline{g}}$ and correlation parameter $\rho_{g\overline{g}} = Corr(\varepsilon_{gi}, \varepsilon_{\overline{g}i})$, and the probability that a consumer picks ethanol is

$$\Pr(i \text{ chooses } e) = \Pr(u_{gi} - u_{ei} \le 0 \cap u_{\overline{g}i} - u_{ei} \le 0)$$

$$= \Phi(-((\alpha(p_{fl}/k_{fi}) + x'_{li}\beta_f) - (\alpha(p_{el}/k_{ei}) + x'_{li}\beta_e)), \Omega_{-e}), \quad f = g, \overline{g},$$

where Φ is the CDF of the bivariate normal random variable $(\varepsilon_g - \varepsilon_e, \varepsilon_{\overline{g}} - \varepsilon_e)$ with mean zero vector and covariance matrix Ω_{-e}.

[23] We reclassify: (i) the 3 premium gasoline consumers as choosing midgrade gasoline; and (ii) the 54 "combo" consumers (see note 17) as choosing the single majority fuel by energy content in the order. Results are robust to dropping, rather than reclassifying, these few observations.

[24] That said, we find that women are three percentage points less likely to choose midgrade gasoline over alternatives than men (not shown; p-value 0.01; $\overline{g}$'s mean choice probability in the sample is 7%).

[25] There is tight correlation between vehicle price and fuel economy (correlation coefficients of –0.64 with k_g and of –0.63 with k_e). That is, pricey cars tend to consume more energy and emit more carbon per km traveled. Appendix A explains how we matched observed vehicles to estimates of their value.

[26] Drivers of expensive vehicles may also be more inclined to adopt an upmarket fuel. We find that such consumers are 2ppt more likely to choose $\overline{g}$, but the effect is not significant (p-value 0.13).

Table 2
Multinomial probit estimated marginal effects on choice of fuel.

Specification	Mean	[I] m.e. (s.e.)	[II] m.e. (s.e.)	[III] m.e. (s.e.)
Prob (Consumer *i* chooses fuel *e*)		Prob = 0.43	Prob = 0.43	Prob = 0.44
Price of *e* (R$/km)	0.27	−3.80*** (0.62)	−4.43*** (0.69)	−6.39*** (0.51)
Price of *g* (R$/km)	0.25	3.27*** (0.63)	4.08*** (0.68)	4.73*** (0.55)
Price of $\tilde{g}$ (R$/km)	0.26	0.53* (0.31)	0.35 (0.27)	1.66*** (0.33)
Female (DV)	0.34	−0.01 (0.03)	−0.01 (0.03)	−0.01 (0.03)
Aged 25–40 years (DV)	0.46	−0.07 (0.04)	−0.06 (0.04)	−0.07* (0.04)
Aged 40–65 years (DV)	0.40	−0.07* (0.04)	−0.06 (0.04)	−0.08** (0.04)
Aged more than 65 years (DV)	0.04	−0.27*** (0.05)	−0.25*** (0.06)	−0.28*** (0.08)
Secondary school (and no more) (DV)	0.31	0.04 (0.05)	0.06 (0.05)	0.04 (0.05)
College educated (DV)	0.62	0.02 (0.05)	0.04 (0.05)	0.03 (0.05)
Extensive vehicle usage (DV)	0.23	−0.08*** (0.03)	−0.09*** (0.03)	−0.09*** (0.03)
Expensive vehicle (DV)	0.26	−0.05** (0.03)	−0.06** (0.03)	−0.05* (0.03)
Value of 12 cars sampled in station (R$m)	0.40	0.28 (0.33)	0.17 (0.35)	0.11 (0.31)
City fixed effect: São Paulo	0.33	0.10 (0.15)	0.11 (0.16)	
City fixed effect: Curitiba	0.22	0.16 (0.14)	0.21 (0.15)	
City fixed effect: Recife	0.11	0.08 (0.14)	0.10 (0.15)	
City fixed effect: Rio de Janeiro	0.11	−0.04 (0.14)	−0.00 (0.15)	
City fixed effect: Belo Horizonte	0.11	−0.14 (0.13)	−0.09 (0.15)	
City fixed effect: Porto Alegre	0.11	−0.23* (0.13)	−0.19 (0.15)	
Stated “Environm.” & $e-$premium≥10% (DV)	0.02		0.51*** (0.03)	
Stated “Environm.” & $g-$premium≥10% (DV)	0.01		0.27* (0.15)	
Stated “Environm.” & similar prices (DV)	0.03		0.37*** (0.06)	
Stated “Engine” & $e-$premium≥10% (DV)	0.07		−0.20*** (0.06)	
Stated “Engine” & $g-$premium≥10% (DV)	0.01		−0.37*** (0.04)	
Stated “Engine” & similar prices (DV)	0.04		−0.24*** (0.05)	
Prob (Consumer *i* chooses fuel *g*)		Prob = 0.50	Prob = 0.51	Prob = 0.48
Price of *e* (R$/km)	0.27	3.27*** (0.63)	4.08*** (0.68)	4.73*** (0.55)
Price of *g* (R$/km)	0.25	−3.87*** (0.84)	−4.74*** (0.95)	−5.47*** (1.10)
Price of $\tilde{g}$ (R$/km)	0.26	0.59 (0.61)	0.66 (0.70)	0.74 (1.23)
(…other marginal effects omitted)				
Prob (Consumer *i* chooses fuel $\tilde{g}$)		Prob = 0.07	Prob = 0.06	Prob = 0.08
Price of *e* (R$/km)	0.27	0.53* (0.31)	0.35 (0.27)	1.66*** (0.33)
Price of *g* (R$/km)	0.25	0.59 (0.61)	0.66 (0.70)	0.74 (1.23)
Price of $\tilde{g}$ (R$/km)	0.26	−1.12 (0.74)	−1.01 (0.80)	−2.40* (1.29)
(…other marginal effects omitted)				
Number of consumers		2160	2160	2160
Total number of alternatives		6288	6288	6288
Log likelihood		−1697	−1618	−1764
$\sigma_{\tilde{g}}$		1.44 (1.08)	2.06 (1.49)	1.37 (1.47)
$\rho_{g\tilde{g}}$ [in III: $\rho_{e\tilde{g}}$]		0.01 (0.83)	0.32 (0.68)	0.56 (0.62)

Note: Estimated marginal effects (m.e.), reported at the sample mean, for baseline specifications. Some effects are not reported due to space constraints. An observation is an alternative that an FFV-qualifying motorist faces among regular gasoline (always available), ethanol (always available) and midgrade gasoline (when available at the station). “DV” denotes a dummy variable. Station visit-clustered standard errors (s.e.) in parentheses.

* $p < 0.1$.

** $p < 0.05$.

*** $p < 0.01$.

less responsive to price. No such evidence came up in the interviews when consumers were open-endedly asked about the basis for their fuel choice.

Rather than nuisance parameters, we interpret the estimated city fixed effects as indicative of home bias. We find that every one of the three capitals of ethanol-producing states displays a stronger taste for ethanol relative to each of the other ethanol-importing cities, and this difference is statistically significant in *all* nine pairwise tests (per specification I). Specifically, against importers Rio de Janeiro, Belo Horizonte and Porto Alegre, *p*-values for equality tests for producer: (i) São Paulo are 0.03, 0.00 and 0.00, respectively; (ii) Curitiba are 0.00, 0.00 and 0.00, respectively; and (iii) Recife are 0.01, 0.00 and 0.00, respectively. This result is consistent with Conjecture 6 (Home Bias). Alternatively, one can interpret these estimates as evidence of long-run habit formation.

Comparing specification II against I, price and other effects are robust to including stated-reason dummies. One may be willing to take the view that a consumer's unassisted and unframed response to the “main reason for having chosen fuel” question, when other than price, is a signal (even if noisy) of the relative weight he places on a non-price characteristic. Alternatively, specification II can be viewed as descriptive, since no stated reason turns out to be a perfect predictor of fuel choice. We include controls for consumers who, when interviewed just after placing their order, spontaneously invoked “the environment" or some aspect (maintenance, performance) of "the engine." We do this by interacting reason dummies with relative price levels, namely dummies for large ethanol price premia (10% or more, i.e., $(p_{el}/k_{ei})/(p_{gl}/k_{gi})\geq 1.1$), large gasoline price premia, and similar relative prices. Our rationale for specifying such interactions is that an environmentally concerned consumer, who perceives ethanol to be "greener" and is willing to pay moderately for this attribute, is more likely to invoke the environment when ethanol is *moderately* more expensive than gasoline (and he stays with ethanol), than when ethanol is way more expensive (in which case he might have switched to gasoline and would not reference the environment) or when ethanol is substantially cheaper (in which case he would no longer need to trade the environment against price and most other consumers would also choose ethanol—recall (1)).[27]

On controlling for stated reasons, we find that invoking the environment is associated with a substantial increase in the probability that the consumer's choice was ethanol over gasoline (unconditionally, in the data, 83% of consumers who referenced the environment had purchased ethanol). We interpret this correlation as being consistent with Conjecture 5 (Environmental Concerns). Consumption of the renewable fuel is associated with a less-deleterious effect on the natural environment. On the other hand, consumers who invoked an engine-related reason are more likely to have chosen gasoline (80% of motorists expressing concern with the engine avoided ethanol). This correlation, while summarized in Conjecture 3 (Technological Concerns), stands at odds with statements by automakers and the specialized press that FFVs are equipped to operate similarly on any blend of gasoline and ethanol (e.g., [21]).

Finally, we investigate how estimates change on dropping city fixed effects—which we interpreted as evidence of home bias (Conjecture 6, Home Bias). In specification III, price effects grow in magnitude, and other effects are robust. Intuitively, price sensitivity is now additionally estimated off of cross-city variation: sugarcane-growing locations exhibit low ethanol prices and strong ethanol adoption relative to ethanol-importing regions, and this correlation is now being picked up by a higher price coefficient α. Typically, concern about price endogeneity in demand studies goes the other way: a strong unobserved taste for a good might correlate positively with prices, as firms take this taste into account on setting prices. In our context, however, that a strong taste for ethanol in sugarcane-growing locations (suggested via the city fixed effects of specification I) correlates negatively with prices suggests a supply side explanation for cross-city price variation. One can plausibly interpret the larger price effect of specification III as a long-run response that also works through changes in long-term habits. We will subsequently show that *even* under this more-price-sensitive specification, fuel switching occurs over a wide range of price variation, highlighting the considerable consumer heterogeneity and departure from perfect substitution.

Alternative specifications: *relative energy*, *and* "*random coefficients*". Table 3 shows that the above results are largely robust to plausible variations in relative energy differences assumed by consumers, as well as to using different subsets of the data. Rather than consider "true" vehicle-specific fuel prices p_{fl}/k_{fi} in R$ per km, specification IV assumes that motorists use the media-reported 70% "conversion rate" to form a price heuristic for ethanol around the per-liter price of gasoline, since the latter has been relatively stable. Thus, facing per-liter pump prices of $(p_{gl}, p_{el}, p_{\overline{g}l})$, a consumer would denominate these in the same "currency" $(0.7p_{gl}, p_{el}, 0.7p_{\overline{g}l})$ irrespective of his FFV. For example, a consumer facing p_{gl} and p_{el} of 2.729 and 2.199 R$/l, respectively, would compare $2.729 \times 0.7 \simeq 1.910$ to 2.199, applying the heuristic $p_{el} < 1.91 \Rightarrow$"*e* priced below *g*." Due to the different units, price effects reported in column IV are not directly comparable to specification I, but they are similar, if slightly larger.[28] This slightly larger price sensitivity suggests that the widely reported 70% parity ratio may be more ingrained than the specific fuel economy ratio for one's FFV—we subsequently offer phone-based interview evidence of this. Other effects are similar.

Another specification we experimented with but do not show for brevity, goes further in terms of cognitive simplification. It assumes that motorists respond only to price comparisons they hear on the local radio, which are overwhelmingly based on median fuel prices surveyed in the city the week before. We follow the radio and use median $(0.7p_g, p_e)$ from the previous week's ANP sample. The similar results we obtain indicate that the moderate relative price dispersion across stations within a city—switched off in this specification—is not significantly driving price sensitivity in our baseline results.[29]

Specification V restricts the sample to the two (ethanol-producing) cities surveyed over multiple weeks—namely the five city-weeks for São Paulo and Curitiba (and prices are again p_{fl}/k_{fi}). Price effects are estimated off of time variation. To test for an "information diffusion" effect, we control for the increase in the ethanol price premium over the preceding fortnight.[30] For perspective, this "recent rate of increase in the relative ethanol price" averages 0.027 R$/km for the early São Paulo

[27] Results are robust to alternative cutoffs (e.g., for the environment dummy, an ethanol price premium of at least 15%, an ethanol premium less than 15%, and any gasoline premium), noting that, empirically, observations are required in each bin. Results are also robust to no interaction. For perspective, with regard to the employed cutoffs, whereas 5.6% of the full sample ($N=2160$) invoked the environment, a lower 4.7% and 4.5% of consumers invoked the environment on facing premia of at least 10% on ethanol ($N=974$) and gasoline ($N=382$), respectively, to be compared with a higher 7.2% of consumers who invoked the environment on facing moderate (mostly ethanol) premia ($N=804$).

[28] Sample means for $(0.7p_{gl}, p_{el}, 0.7p_{\overline{g}l})$ are (1.88,1.76,1.83) R$/l (omitted for brevity).

[29] Price effects under this "newscast" specification are slightly higher than in column IV, offering suggestive evidence that some consumers complement prices they read at the pump with more salient information reported on the local radio. This robustness test—and others that are not reported, e.g., controlling for the number of nozzles dispensing each fuel at the station—are available upon request.

[30] We include the weakly positive covariate $\max\{0, (p_e - L2(p_e))/k_{ei} - (p_g - L2(p_g))/k_{gi}\}$, with $L2$ denoting the two-week lag operator over the city's median price. We also note that in both specification I and this one, which exploit time-series variation, clustering on station (rather than on station-visit) to account for persistent unobserved shocks yields fairly similar, if lower, standard errors.

Table 3
Multinomial probit estimated marginal effects on choice of fuel.

Robustness (energy differences/"random coefficients")	"Media parity" [IV] m.e. (s.e.)	"2 cities, time" [V] m.e. (s.e.)	"Extensive user" [VIa] m.e. (s.e.)	"Lighter user" [VIb] m.e. (s.e.)	"Higher income" [VIIa] m.e. (s.e.)	"Lower income" [VIIb] m.e. (s.e.)	"Price diff. $< 10\%$ " [VIII] m.e. (s.e.)
Prob (Consumer *i* chooses fuel *e*)	Prob = 0.43	Prob = 0.60	Prob = 0.41	Prob= 0.47	Prob = 0.41	Prob = 0.45	Prob = 0.54
Price of e (IV. R$/l, V on. R$/km)	−0.59*** (0.09)	−4.83*** (0.99)	−4.45*** (0.78)	−2.69** (1.06)	−3.57*** (0.71)	−4.20*** (0.96)	−6.81*** (2.11)
Price of g (IV. R$/l, V on. R$/km)	0.53*** (0.09)	4.54*** (0.81)	4.41*** (0.74)	1.99 (1.43)	3.22*** (0.81)	3.68*** (0.88)	6.17*** (2.24)
Price of $\overline{g}$ (IV. R$/l, V on. R$/km)	0.06 (0.05)	0.29 (0.88)	0.04 (0.14)	0.71 (0.82)	0.35 (0.61)	0.52 (0.49)	0.64 (0.62)
Female (DV)	−0.01 (0.03)	−0.01 (0.06)	0.02 (0.05)	−0.01 (0.06)	0.00 (0.04)	−0.02 (0.04)	0.02 (0.07)
Aged 25–40 years (DV)	−0.07* (0.04)	−0.12** (0.06)	−0.08 (0.08)	−0.03 (0.06)	−0.13** (0.06)	−0.02 (0.05)	−0.07 (0.09)
Aged 40–65 years (DV)	−0.07* (0.04)	−0.10* (0.06)	−0.07 (0.06)	−0.04 (0.08)	−0.14** (0.06)	−0.00 (0.05)	−0.09 (0.11)
Aged more than 65 years (DV)	−0.26*** (0.05)	−0.29* (0.15)	−0.28** (0.11)	−0.25**(0.10)	−0.29*** (0.07)	−0.22*** (0.09)	−0.31** (0.15)
Secondary school (and no more) (DV)	0.04 (0.05)	0.07 (0.07)	0.09 (0.08)	0.05 (0.08)	0.01 (0.08)	0.08 (0.07)	0.10 (0.08)
College educated (DV)	0.02 (0.05)	0.05 (0.05)	0.03 (0.08)	0.07 (0.08)	0.02 (0.07)	0.03 (0.07)	0.03 (0.08)
Extensive vehicle usage (DV)	−0.08*** (0.03)	−0.08* (0.05)			−0.08* (0.04)	−0.08* (0.04)	−0.10** (0.04)
Expensive vehicle (DV)	−0.06** (0.03)	−0.11*** (0.04)	−0.05 (0.04)	−0.10 (0.05)			−0.07 (0.05)
Value of 12 cars sampled in station (R$m)	0.20 (0.32)	0.59 (0.62)	−0.03 (0.43)	0.54 (0.50)	−0.12 (0.40)	0.54 (0.46)	0.42 (0.53)
City fixed effect: São Paulo	0.10 (0.16)	0.07 (0.44)	0.13 (0.48)	−0.07 (0.23)	0.29 (0.23)	−0.07 (0.22)	0.30 (0.26)
City fixed effect: Curitiba	0.16 (0.15)	0.07 (0.28)	0.17 (0.40)	0.06 (0.22)	0.31 (0.21)	0.06 (0.22)	0.10 (0.26)
City fixed effect: Recife	0.10 (0.14)		0.14 (0.26)	−0.10 (0.21)	0.23 (0.20)	−0.03 (0.20)	−0.03 (0.25)
City fixed effect: Rio de Janeiro	−0.01 (0.15)		0.02 (0.28)	−0.21 (0.22)	0.13 (0.23)	−0.16 (0.19)	−0.24 (0.21)
City fixed effect: Belo Horizonte	−0.10 (0.14)		−0.03 (0.29)	−0.39 (0.19)	0.07 (0.22)	−0.29 (0.16)	
City fixed effect: Porto Alegre	−0.22 (0.15)		−0.23 (0.23)	−0.36 (0.16)	−0.09 (0.25)	−0.34 (0.18)	
Past 2-week rise in rel. price of e (R$/km)		2.28 (2.39)					
Prob (Consumer *i* chooses fuel *g*)	Prob = 0.50	Prob = 0.36	Prob = 0.52	Prob= 0.46	Prob = 0.50	Prob = 0.50	Prob =0.43
Price of e (IV. R$/l, V on. R$/km)	0.53*** (0.09)	4.54*** (0.81)	4.41*** (0.74)	1.99 (1.43)	3.22*** (0.81)	3.68*** (0.88)	6.17*** (2.24)
Price of g (IV. R$/l, V on. R$/km)	−0.64*** (0.16)	−5.26* (2.71)	−5.66*** (1.55)	−3.14*** (0.94)	−4.44*** (1.66)	−4.10*** (0.97)	−6.33*** (2.27)
Price of $\overline{g}$ (IV. R$/l, V on. R$/km)	0.10 (0.13)	0.72 (2.31)	1.24 (1.29)	1.15 (1.50)	1.21 (1.30)	0.42 (0.43)	0.17 (0.21)
Number of consumers	2160	1200	1063	772	1100	1060	804
Total number of alternatives	6288	3528	3112	2240	3214	3074	2336
Log likelihood	−1691	−917	−835	−604	−888	−796	−640
$\sigma_{\overline{g}}$	1.78 (1.22)	1.81 (0.93)	2.67 (1.81)	0.43 (1.61)	1.58 (1.31)	1.51 (1.59)	1.90 (2.26)
$\rho_{g,\overline{g}}$	0.27 (0.94)	0.66 (2.04)	0.97 (0.24)	0.74 (2.30)	0.64 (0.76)	−0.24 (1.02)	−0.73 (1.28)

Note: Estimated marginal effects, reported at the sample mean, for alternative specifications. "DV" denotes a dummy variable. Some effects are not reported (see text and the notes to Table 2). Station visit-clustered standard errors in parentheses.

* $p < 0.1$.

** $p < 0.05$.

*** $p < 0.01$.

Table 4
Probit estimated marginal effects on choice of expensive fuel.

Specification	Mean	[1] m.e. (s.e.)	[2] m.e. (s.e.)	[3] m.e. (s.e.)	[4] m.e. (s.e.)	[5] m.e. (s.e.)
Female (DV)	0.36	−0.02 (0.03)	−0.02 (0.03)	−0.01 (0.03)	−0.01 (0.03)	−0.02 (0.03)
Aged 25–40 years (DV)	0.46	0.08 (0.05)	0.07 (0.05)	0.07 (0.05)	0.06 (0.05)	0.09* (0.05)
Aged 40–65 years (DV)	0.40	0.05 (0.05)	0.03 (0.06)	0.03 (0.06)	0.02 (0.05)	0.07 (0.05)
Aged more than 65 years (DV)	0.04	0.04 (0.10)	0.02 (0.10)	0.03 (0.10)	0.01 (0.10)	0.09 (0.12)
Secondary school (and no more) (DV)	0.33	0.06 (0.06)	0.05 (0.06)	0.05 (0.06)	0.07 (0.06)	0.03 (0.06)
College educated (DV)	0.60	0.04 (0.06)	0.01 (0.06)	0.02 (0.06)	0.01 (0.06)	−0.01 (0.06)
Expensive vehicle (DV)	0.28	0.03 (0.04)	0.02 (0.04)	0.02 (0.04)	0.02 (0.04)	0.05 (0.03)
Value of 12 cars in station (R$m)	0.41	−0.12 (0.37)	−0.03 (0.39)	−0.08 (0.38)	0.07 (0.37)	−0.13 (0.35)
km purchased on occasion (× 100)	1.90	−0.08*** (0.02)	−0.08*** (0.02)	−0.07*** (0.02)	−0.06*** (0.02)	−0.05*** (0.02)
Stated “I chose out of habit” (DV)	0.03	0.42*** (0.08)	0.41*** (0.08)	0.43*** (0.08)	0.38*** (0.09)	0.30*** (0.10)
Price premium e above g	0.04	−4.11*** (1.16)	0.26 (1.53)			
Price discount e below g	0.01	−12.41*** (1.65)	−7.53*** (1.83)			
2-week variation in relative prices	0.01	10.18*** (2.51)	5.51** (2.78)			
City fixed effect: São Paulo	0.26		−0.04			
City fixed effect: Curitiba	0.24		−0.07			
City fixed effect: Recife	0.05		0.12	0.23	0.14	0.13
City fixed effect: Rio de Janeiro	0.13		−0.12	−0.05	−0.11	−0.09
City fixed effect: Belo Horizonte	0.19		−0.20*	−0.16	−0.21**	−0.17
City fixed effect: Porto Alegre	0.12		−0.26***	−0.25***	−0.27***	−0.25***
City-week FE: São Paulo, January 11	0.05			0.10	0.01	0.10
City-week FE: São Paulo, January 25	0.07			−0.05	−0.12	−0.13
City-week FE: São Paulo, March 29	0.14			−0.11	−0.18	−0.16
City-week FE: Curitiba, January 25	0.09			0.07	−0.01	0.11
City-week FE: Curitiba, March 29	0.15			−0.21**	−0.24***	−0.20**
Number of observations		1002	1002	1002	1011	1003
Log likelihood		−525	−511	−508	−517	−503

Note: Estimated marginal effects, reported at the sample mean (reported means correspond to the covariates in specifications 1–4). Success is “Expensive fuel is chosen over the cheaper substitute.” An observation is a consumer in the restricted sample facing sufficiently unequal prices across regular gasoline and ethanol. See the text for price variable definitions and units. “DV” and “FE” denote dummy variable and fixed effect, respectively. Station visit-clustered standard errors in parentheses. For city and city-week fixed effects, only significance levels are shown due to space.

* $p < 0.1$.

** $p < 0.05$.

*** $p < 0.01$.

January 11 subsample, immediately following the ethanol price hike, compared to less than 0.001 R$/km for the São Paulo January 25 subsample, as prices had plateaued. While the marginal effect on the probability of choosing ethanol is, as expected, positive—suggestive of gradual rather than instant information diffusion among some consumers—it is not significant, neither statistically nor economically. Indeed, in the raw data, the proportion of São Paulo consumers paying an average 8% premium for ethanol was 51% in early January compared to a rather similar 42% a fortnight later. We return to information diffusion below. Importantly, price and non-price effects are similar to specification I. Specification V is akin to allowing “random coefficients” for the São Paulo and Curitiba subsample relative to the other cities.

The columns marked “VIa” and “VIb” report marginal effects obtained from separately fitting specification I on two partitions of the data based on stated vehicle usage: the 1063 motorists at or above the median of 200 km/week, and the 772 motorists with below-median usage (325 motorists stated not knowing their usage). This specification can be viewed as very flexibly introducing random coefficients on price and non-price characteristics in (2), based on stated vehicle usage. The price sensitivity for consumers who profess to drive their vehicle more is estimated to be somewhat larger than that of below-median commuters, but, importantly, substitution across fuels still takes place over a wide range of relative prices—see below. Non-price effects across the subsamples are similar, including age and the pairwise “home bias” test results.

Similarly, we separately estimate specification I on two partitions of the data based on vehicle price, as a proxy for a consumer's socioeconomic standing—at or above the median of R$ 29,504 versus below the median. As shown in columns “VIIa” and “VIIb,” price and non-price effects are similar to specification I; if anything, drivers of cheaper vehicles appear slightly more price sensitive. We also estimate specification I using only a subset of choices for which fuel price differences were moderate, i.e., we drop observations where consumers faced either an ethanol or a gasoline price premium of at least 10%—see column VIII. Estimates are similar but less precise. Price effects, now estimated off of this subsample's narrower variation (e.g., there are no Belo Horizonte or Porto Alegre markets), are comparable if somewhat larger than some of the earlier estimates. Overall, results suggest that our baseline specification is not excessively restrictive and that effects are identified from across the empirical distribution.[31]

[31] Other robustness tests include estimating specification I separately on two equally sized partitions of the data based on our proxy for the average income of a station's shoppers: observations collected in above-median- versus below-median-income station visits, where the value of the 12-vehicle sample averages R$ 0.37 m and R$ 0.44 m per station, respectively. Estimated effects are similar.

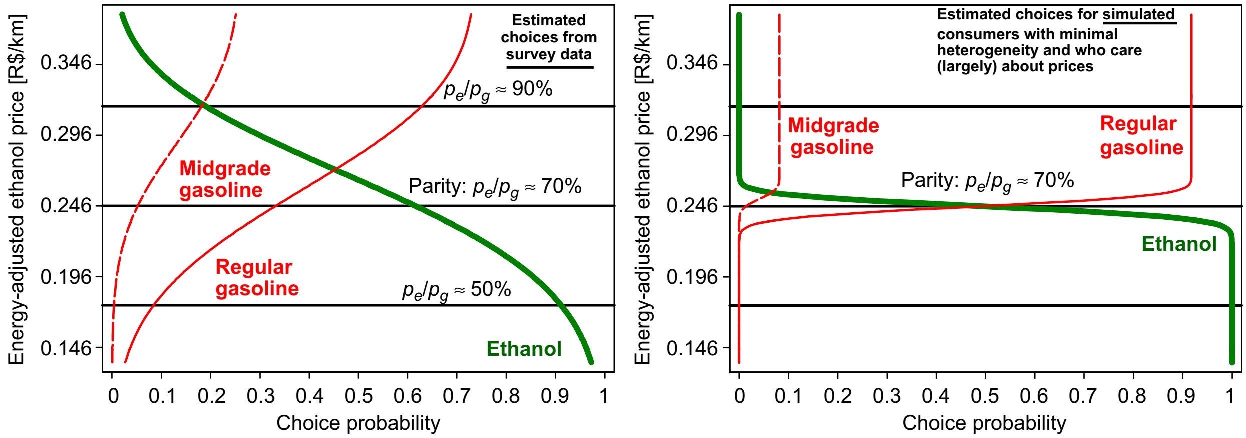

Fig. 6. Illustrating unobserved consumer heterogeneity. Left panel: A "median" consumer's fuel choice probabilities estimated from the survey data. Source: Specification III estimates, to conservatively reduce the price range over which substitution takes place. Right panel: Fuel choices probabilities estimated for (2160) simulated consumers when unobserved heterogeneity is small and mostly prices matter. The panels indicate fuel choice probabilities as the energy-adjusted price of ethanol, in R$/km, is varied while holding gasoline prices constant at the sample means (regular 0.246 R$/km, midgrade 0.256 R$/km) and preserving three fuels in the consumer's choice set. The equivalence scale with respect to the ethanol-to-regular-gasoline per-liter price ratio is indicated by the horizontal lines (for an ethanol-to-gasoline fuel economy ratio of 70%).

Illustrating consumer heterogeneity. What choices does an "average" consumer make at different relative prices? Consider a male motorist aged 25–40 years, who states having at least some college education, who neither states driving extensively nor drives an expensive vehicle. In the left panel of Fig. 6, we plot choice probabilities for this median consumer using specification III (Table 2) estimates, for which price sensitivity was estimated to be highest. We employ this specification, which exploited cross-city variation by dropping city fixed effects, rather than specification I, which kept them, as our intention is to conservatively reduce the range of price variation over which fuel switching takes place: as we show, substitution still occurs over a wide range of relative price variation, not only around parity. To produce the figure, we vary the per-km ethanol price while holding, in a choice set $\{g, e, \overline{g}\}$, the prices of regular and midgrade gasoline constant at their sample means of 0.246 R$/km and 0.256 R$/km, respectively.

As the left panel indicates, when e is priced at parity to g (i.e., 0.246 R$/km), the probability of choosing e is just over 60%, and the choice probabilities for g and $\overline{g}$ are just under 35% and 5%, respectively. What is striking is that even for this median consumer, when e is priced at a substantial premium relative to g—say $p_e/k_e = 0.316$ R$/km, equivalent to a 29% premium over g—the choice probability for e is still a sizable 19%! Similarly, when e is priced at a substantial discount relative to g—say a 29% discount—the probability that this consumer still buys gasoline (regular or midgrade) is a non-negligible 9%.[32] This heterogeneous response to prices among consumers with a given set of observed characteristics is evidence of considerable unobserved heterogeneity. To draw a contrast with survey data, the right panel of Fig. 6 plots choice probabilities estimated off a sample of 2160 consumers facing actual prices and choice sets but all of whom we *simulate* to care mostly about prices. That is, for this simulation, in utility function (2) we *set*: (i) non-price coefficients on observables at zero, $\beta = 0$; (ii) a large relative price coefficient, $\alpha/\sigma = 200$; and (iii) a covariance matrix Ω containing $\sigma = 1$ along the diagonal and $\rho = 0$ off it.

As for observed characteristics, we illustrate the similarly extensive consumer heterogeneity in Fig. 7, in two ways. In the left panel, now employing specification I estimates, with city fixed effects, we plot the ethanol choice probability for each of two hypothetical polar consumers: (i) an "ethanol fan," defined as a young male aged 25 years– with some college education and who resides in Curitiba, the capital of sugarcane-growing Paraná state; and (ii) a "gasoline fan," defined as an older male aged 65 years+ with no more than primary education, who drives an expensive vehicle and drives it extensively, and resides in ethanol-importing Porto Alegre. The difference is stark. The polar types illustrate the wide range of variation in behavior.

A second way by which to illustrate consumer heterogeneity is to compare estimates obtained from separate subsets of the data. The right panel of Fig. 7 plots ethanol choice probabilities for consumers with above-median versus below-median stated vehicle usage, as in columns "VIa" and "VIb" of Table 3 (the plot considers otherwise median consumers in Rio de Janeiro). Relative to less extensive users, motorists who drive more extensively are somewhat more price sensitive and, faced with price parity, tend to favor gasoline over ethanol.

Another test that consumers respond to more than price is to replace the city fixed effects in specification I with city-week fixed effects, further soaking up temporal price variation (one can go even further and specify station-visit fixed effects

[32] Had we based Fig. 6 on the less-price-sensitive specification I, with the Rio fixed effect, the larger price range over which switching occurs would be symmetric about parity—see Fig. 8.

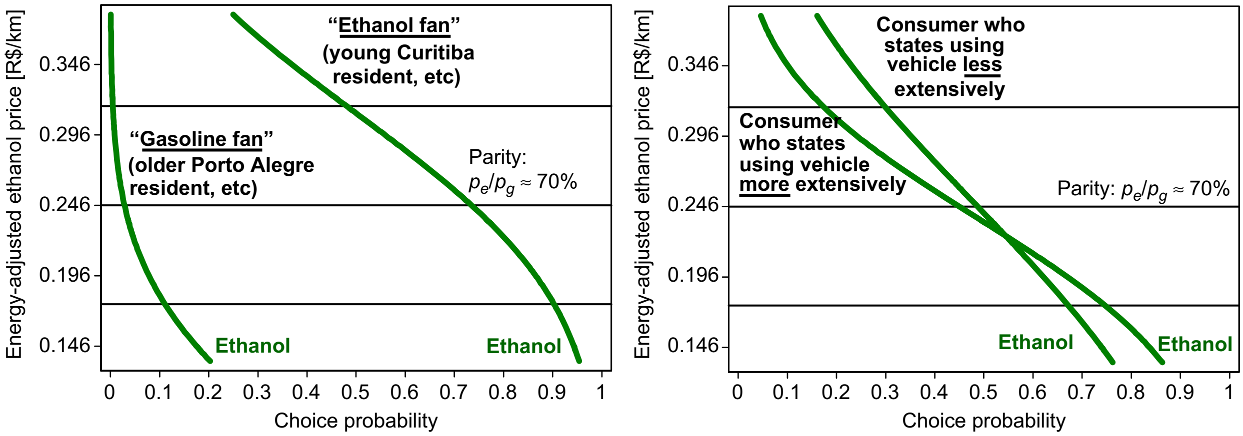

Fig. 7. Illustrating observed consumer heterogeneity. Left panel: Ethanol choice probabilities for two hypothetical polar consumers, an "ethanol fan" and a "gasoline fan." Source: Specification I estimates. Right panel: Ethanol choice probabilities for consumers with above-median versus below-median stated vehicle usage. Source: Specification VIa and VIb estimates. The panels indicate ethanol choice probabilities as the energy-adjusted price of ethanol, in R$/km, is varied while holding gasoline prices constant at the sample means and preserving three fuels in the consumer's choice set. The equivalence scale with respect to the ethanol-to-regular-gasoline per-liter price ratio is indicated by the horizontal lines (for an ethanol-to-gasoline fuel economy ratio of 70%).

to throw out price dispersion across stations). We find very robust (non-price) effects, such as older consumers or extensive vehicle users favoring gasoline, including magnitude.[33]

"*Willingness to pay for greenness*." We momentarily consider specification II estimates, which include stated-reason dummies, and ask, how much more (resp., less) likely is a consumer spontaneously invoking the environment to have purchased ethanol (resp., gasoline)? Fig. 8 plots choice probabilities for median consumers, as just defined, in each of three cities exhibiting a decreasing degree of ethanol home bias: Curitiba, Rio de Janeiro and Porto Alegre. Ethanol and gasoline (regular or midgrade) choice probabilities are plotted in the left and right panels, respectively, again as ethanol prices are varied while holding gasoline prices and the three-alternative choice set constant. The vertical shifts depict the effect of switching the main-reason-was-environment dummy from "off" to "on," marked by the thin and thick lines respectively.[34] To illustrate, an environment-invoking Rio consumer facing an ethanol price of 0.34 R$/km has the same 50% probability of adopting ethanol as a non-environment-invoking Rio consumer facing a substantially lower ethanol price of 0.25 R$/km. One may interpret this 0.09 R$/km (0.08 US $/mile) shift as a measure of the "willingness to pay for greenness" by a subset of the population.

5. Follow-on phone interviews

In late 2010 we hired a market research firm that specializes in telephone marketing to contact consumers in our sample for additional questions. Of the 2160 motorists observed at the station, 1991 had provided contact details at the end of their interview. An interviewer worked on the 1991 contacts for over a month and managed to conclude 607 phone-based interviews.

After reminding the consumer about the face-to-face interview earlier that year, the first question asked "In your view, the consumption of which motor fuel pollutes less and is better for the environment: gasoline, ethanol or is there no difference?" We randomly assigned the ordering of the fuels in the statement, to control for potential framing. An overwhelming 82%, or 500 consumers, selected ethanol, and a further 13% felt there was no differential impact on the environment (recall the motivation behind Conjecture 5, Environmental Concerns, in Table 1).

The next question used similar words to address any differential impact of the two fuels on "the performance of your flex car": 49% of telephone respondents perceived gasoline to yield better performance, against 31% stating ethanol and 20% perceiving no difference across the two fuels (Conjecture 3, Technological Concerns).

A third question attempted to assess home bias, asking whether the consumption of either fuel was "better for Brazil," to which 70%, or 422 consumers, responded with a favorable view of ethanol, and a further 25% said ethanol or gasoline consumption had the same impact on their country. Those 422 consumers who stated that Brazil benefited from substituting ethanol for gasoline were then simply asked "Why?": 194 consumers (46%) spontaneously justified their answer by reference to some combination of (ethanol) "is a local product," "employs local technology," and/or "creates local jobs."

It is remarkable how favorable views of ethanol over gasoline—on the environment, vehicle, or country—correlate with whether the respondent resides in a sugarcane-growing state (Conjecture 6, Home Bias). For example, that ethanol

[33] We thank a reviewer for suggesting this test.

[34] We illustrate by turning on the interaction "Stated 'Environment' & similar prices" over the entire price range. Had we turned on each of the three price interactions at selected price points, the adoption curves would jump at the cutoffs, but the shifts would not change significantly.

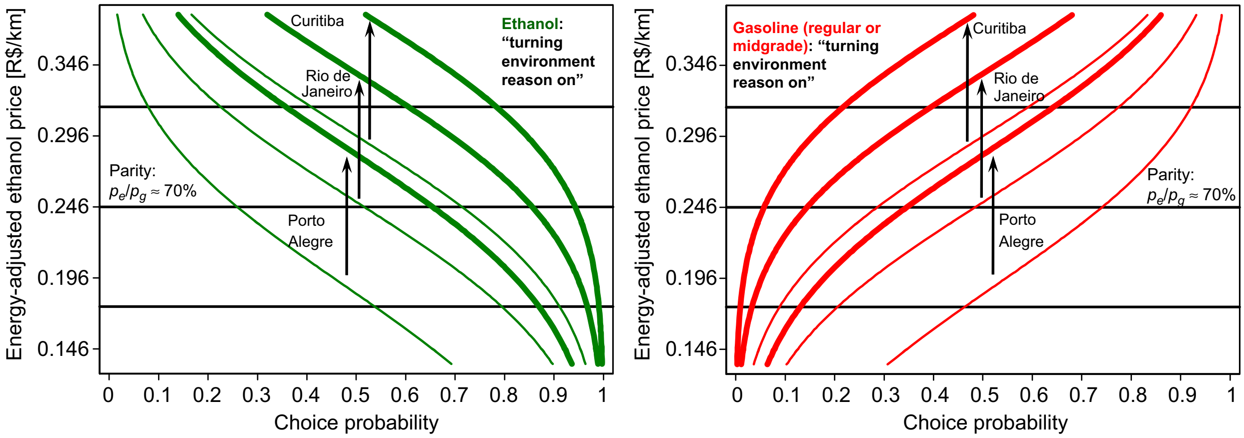

Fig. 8. "Willingness to pay for greenness." Choice probabilities for ethanol (left panel) and for gasoline (both varieties, right panel), when the "main reason was the environment" dummy is switched from off (thin lines) to on (thick lines). Source: Specification II estimates. The plots consider a "median" consumer in each of three cities, as the energy-adjusted price of ethanol, in R$/km, is varied holding gasoline prices constant at the sample means and preserving three fuels in the consumer's choice set. The equivalence scale with respect to the ethanol-to-regular-gasoline per-liter price ratio is indicated by the horizontal lines (for an ethanol-to-gasoline fuel economy ratio of 70%).

consumption "is better for Brazil" was the opinion of 74% of residents in São Paulo, Curitiba and Recife (respectively, 75%, 74% and 69%; $N'_{producer} = 416$) against a statistically significantly lower 61% of residents in ethanol-importing locations who felt the same (Rio de Janeiro 64%, Belo Horizonte 58% and Porto Alegre 59%; $N'_{importer} = 191$; one-tailed p-value of 0.001). In a similar vein, 55% of respondents in "pro-gasoline" Porto Alegre felt their FFV got more performance out of gasoline against a significantly lower 38% of respondents in "pro-ethanol" Curitiba.[35] See Appendix B for all the by-city proportions across the three questions.[36]

The follow-on phone interviews also probed into consumers' understanding of price parity across ethanol and gasoline, whether they applied the parity concept by calculating p_{fi}/k_{fi} in R$/km, compared the per-liter p_{el}/p_{gl} ratio against the media-reported 70% rate, or formed the gasoline-based price heuristic for ethanol $0.7p_{gl}$. All 607 phone participants were asked "How do you decide on which fuel to purchase?"; the 403 consumers whose response was clearly based on price[37] were then further asked "How do you calculate whether ethanol is favorably or unfavorably priced?" Over 60% of these 403 respondents were able to reasonably describe parity or allude to the different energy content of the fuels. The remaining 40% either appeared confused, stated that they did not know or do the cost conversion themselves (relying instead on a relative, the serviceman, or the media/radio), or skirted the question. (See Appendix B for details.)

We find the phone-based evidence reassuring in light of our earlier interpretations.

6. Policy implications

Our results offer important policy implications for energy planners. The imperfect demand substitutability that we find suggests that policies to raise the share of alternative fuels should focus as much on the demand side as on the supply side. For example, former US Senator Richard G. Lugar, writing in the Wall Street Journal in 2006 [17], seemed to take consumer behavior for granted: "(s)witching to an ethanol-based transportation system, by adapting new cars to run on an ethanol–gasoline blend with inexpensive, off-the-shelf flexible fuel technology and piggy-backing on the existing gas station network, would be both good policy and a great bargain for the American consumer." Policymakers may seek to change relative prices, but as our estimates suggest, the demand response is likely to be limited insofar as consumers have heterogeneous preferences for the different non-price characteristics—real or perceived—of substitute fuels. The price counterfactual analysis we provide next forcefully illustrates this point. Further, aware that consumers form tastes over fuel characteristics other than prices, policymakers could try to steer demand one way or another by addressing such non-price characteristics. This point is underscored by a recent advertising campaign that UNICA, a Brazilian sugar industry trade association, launched in late 2012 after four years without advertising [1]. While making only occasional reference to prices,

[35] Recall Fig. 7. Anecdotally, Curitiba has been labeled Brazil's "green capital," e.g., US magazine Grist [13] places Curitiba third in a list of "15 Green Cities," after Reykjavik, Iceland and Portland, Oregon, US. In response to our telephone survey's first question, 85% of respondents in Curitiba perceived ethanol to be more environmentally friendly than gasoline, compared with a significantly lower 71% among Porto Alegre respondents (p-value 0.01).

[36] Two comments are in order: (i) we find little evidence of selection bias in the subsample of 607 telephone respondents relative to the full sample of 2160 motorists, e.g., women are sampled somewhat less; and (ii) adding "observables" constructed from a consumer's phone responses increases the power of the baseline demand specifications (Table 2) in predicting his fuel choice months earlier, e.g., phone respondents who stated that their vehicle performed better on ethanol were 14–17ppt (depending on the specification) more likely to have purchased ethanol at the station (1% significance level).

[37] We did not enquire about parity if a consumer, e.g., voiced a strong preference for one fuel, such as the 63 consumers who stated (something resembling) "I always choose gasoline/I dislike ethanol."

TV ads, featuring local actors interviewing consumers at the pump, pitch ethanol as the "green" fuel, the "home" fuel, and the fuel that "boosts your vehicle's power."[38]

A relative price counterfactual: *Planning the energy mix*. The estimated demand model can be used to predict market-level shares for the substitute fuels under different relative price scenarios (also see Appendix C). Consider a planner in the Amazonian state of Pará, in the remote north of Brazil, home to 7.6 m people in 2010, two thirds of whom lived in urban areas. Thanks to both a high state sales tax on ethanol and the federal government's favorable gasoline pricing policy toward remote states, Pará state's FFV motorists have grown used to less-expensive gasoline relative to ethanol at the pump: the last time the per-liter p_e/p_g ratio dipped below 70% was in 2002, and since then through May 2010 it hovered between 75% and 90%, with a mean of 82%.[39] In early 2010, FFVs accounted for 45% of Pará's car fleet, thanks to booming new car sales (overwhelmingly FFVs) in recent years.

Suppose the state planner is considering a reduction in the state sales tax ("ICMS") on ethanol, to shift the source of energy powering the state's FFVs to ethanol, away from gasoline. (We are not advocating a particular policy but illustrating the planner's problem.) In May 2010, Pará's nominal ICMS tax rate on ethanol stood at 28%, the highest in the country, to be compared against 12% and 18% in the sugar-producing states of São Paulo and Paraná, respectively. One might think that lowering the pump price of ethanol to equal that of gasoline on an energy-adjusted basis might suffice.

But "will they come?" Suppose that Pará consumers were to display the tastes and behavior of Porto Alegre, the city in our sample where ethanol prices were highest. Our demand estimates suggest that the uptake of ethanol among Pará FFV motorists, were fuels to be priced equally in R$/km, would increase only slightly, from 13% of gasoline-plus-ethanol vehicle km traveled to 19%—see the table below. For illustrative purposes, our analysis keeps the tax calculation simple, ignores a supply response from a lower tax—e.g., ethanol producers raising prices—and assumes the demographic distribution and fuel availability in our survey with the Porto Alegre fixed effect turned on.

Scenario for Pará state	May 2010 "Current"	Counterfactual 1 Price parity	Counterfactual 2 12% ICMS tax (SP)	Counterfactual 3 0% ICMS tax
p_g, R$/l	2.695	2.695	2.695	2.695
p_e, R$/l	2.075	1.887	1.698	1.494
(Ratio) p_e/p_g	77%	70%	63%	55%
ICMS in p_e, R$/l	0.581	0.393	0.204	0
$p_{\overline{g}}$, R$/l (91% avail.)	2.799	2.799	2.799	2.799
Predicted ethanol share of gasoline-plus-ethanol "vehicle km purchased" (Porto Alegre tastes):				
FFVs only, Pará	13%	19%	26%	36%

Note: Pump prices are inclusive of ICMS sales tax. Predictions as per specification I.

Were Pará's state sales tax on ethanol to be lowered to the 12% rate in São Paulo, our predicted ethanol share would rise to only 26%, despite p_e/p_g falling below the 70% parity threshold. It is striking that even if the tax rate were reduced to zero, with ethanol favorably priced at $p_e/p_g = 55\%$, ethanol's share of the FFV energy mix would not reach 50%. Of course, the analysis assumes away any changes in consumer preferences, behavior or information. For example, were Pará's FFV motorists to behave like their counterparts in sugar-producing São Paulo, ethanol adoption on facing a very favorable ethanol price of $p_e/p_g = 55\%$ would rise to 73%. The estimated demand system can be used to perform similar "non-price" counterfactuals, such as raising the proportion of "green-minded" consumers or getting older consumers to behave more like younger ones (particularly if one has estimates of the effectiveness of informational campaigns).

7. Choosing the more expensive fuel

We further examine heterogeneity by considering the choices of the subset of consumers who faced sufficiently unequal energy-adjusted prices across the two standard fuels g and e. In such markets, which observable characteristics help explain the choice of the more expensive fuel over a cheaper substitute?

Our purpose in the following analysis is to look, in a simple way, for suggestive evidence of consumer "inattention" or "inertia" where we are more likely to find it. We discard observations in which consumers faced similarly priced g and e. Further, we simply restrict the analysis to consumers who appear less likely to have a strong taste for one fuel; thus, the

[38] See http://www.etanolverde.com.br/galeria.php?tipo=2. In one copy, a consumer opines that "ethanol's return to the planet is priceless, and pollution is practically zero," while the interviewer, actress D. Winits, states "absolutely, this guy knows everything about ethanol…it is a genuinely Brazilian fuel," and the consumer goes on to say "if everyone picked ethanol they'd contribute to our country, not only in terms of activity but also taxes … (as the fuel) is not imported (but) 100% Brazilian." In another ad, actor R. Sant'Anna asks a consumer: "Did you know that ethanol employs more, many more workers than other fuels?"

[39] That is, were gasoline and ethanol considered perfect substitutes by Pará's FFV motorists, its fueling stations would not have sold ethanol to them even before the early 2010 price hike. But fuel distributor reports (see Appendix C) suggest that even this pricey location has its share of "ethanol fans." For example, ethanol's share of Pará's gasoline-plus-ethanol "vehicle km purchased"—across both FFVs and single-fuel, essentially gasoline-dedicated vehicles —was reported to be 4% in December 2009, with p_e/p_g at 78%. We estimate the penetration of ethanol-dedicated vehicles at only 1% of the state's fleet, thus the bulk of the ethanol share is likely due to demand from FFVs.

descriptive evidence we present next is intended to motivate future research and should not be interpreted structurally. Specifically, we consider consumers who invoked neither environmental nor engine concerns as the basis for their observed fuel choice, *and* who stayed clear of non-standard fuel $\overline{g}$. Consumers who spontaneously expressed concern for a non-price characteristic, or who revealed a preference for $\overline{g}$ over cheaper g, may be more likely to value a characteristic other than price. The table below describes the resulting restricted set of 1002 observations and, in the final column, the proportion of these 1002 consumers who chose the expensive fuel.

City-week, mean p_e/p_g (%)	Number of observations:				Proportion choosing the expensive fuel
	Consumers observed	(−) Facing similar prices	(−) Likely strong taste	(=) Included in analysis	
S.Paulo-Jan11, 74.2	240	175	14	51	26/51 = 51%
S.Paulo-Jan25, 74.7	240	151	17	72	27/72 = 38%
Curitiba-Jan25, 74.7	240	145	8	87	44/87 = 51%
Recife-Jan25, 74.5	240	174	15	51	33/51 = 65%
R.Janeiro-Jan25, 80.5	240	32	74	134	44/134 = 33%
B.Horiz.-Jan25, 84.6	240	0	46	194	41/194 = 21%
P.Alegre-Jan25, 90.5	240	0	116	124	10/124 = 8%
S.Paulo-Mar29, 59.1	240	81	20	139	36/139 = 26%
Curitiba-Mar29, 57.8	240	57	33	150	19/150 = 13%
Total	2160	815	343	1002	280/1002 = 28%

Note: Observations facing "similar" prices, as reflected in the table and dropped in specifications 1–3 of Table 4, are those for which $|p_{gl}/k_{gi} - p_{el}/k_{ei}| < 0.0242$ R$/km, where 0.0242 R$/km is calculated as one-tenth the mean price of the cheaper fuel, $\min\{p_{gl}/k_{gi}, p_{el}/k_{ei}\}$, taking the mean over all 2160 observations in the sample.

Before reporting probit estimates of the choice between the relatively expensive fuel and the cheap alternative, we note two intuitive patterns in the data that will inform our probits. First, ethanol prices in São Paulo on the week of January 11 had already risen and were similar to prices on the week of January 25, yet the proportion of consumers choosing this expensive fuel two weeks later was 13ppt lower. This *difference* is likely due to (short-run) consumer "inattention" rather than "tastes:" there may be a 13% fraction of late switchers who might have switched to g, the cheaper substitute, by January 11 but required an additional fortnight to process the relative price change that had already taken place. As for the *level*—the 38% of São Paulo motorists who were sticking to ethanol despite the 3–4 weeks of high prices—these choices are likely driven by "tastes," which has been the focus of this paper, as well as by (longer-run) "inattention." Second, the proportion of consumers choosing the expensive fuel declines with the magnitude of the price difference: compare the adoption of substantially more expensive ethanol (P.Alegre-Jan25) or substantially more expensive gasoline (Curitiba-Mar29) against the adoption of a moderately more expensive fuel (e.g., Curitiba-Jan25). This pattern can be explained by either "tastes" or "inattention."

Columns 1–3 of Table 4 report marginal effects of probit specifications estimated off the 1002 observations, where the response variable takes on the value 1 when a consumer chooses the expensive fuel (higher p_{fl}/k_{fi}) over the cheaper alternative. The more expensive fuel is g in 289 observations, all collected in March, and e in 713 observations. Controls for gender, age, education and vehicle price are insignificant. The result that educational category is not significant may owe to the inclusive definition of a college degree (Appendix A), and also to the hypothesis that unobserved strong tastes rather than price misconception are driving the choice of the fuel that happens to be more expensive in the given market. Translating each purchase into vehicle km (l × km/l), we find that every additional 100 km purchased is associated with an 8% lower probability that the more expensive fuel is chosen. It is conceivable that consumers with greater "stakes" in the purchase are better informed about effective prices, or they may place less weight on non-price characteristics. We find that consumers who are conscious of "habit" playing a role in their choice at the pump are significantly more likely to pick the expensive fuel relative to consumers who do not spontaneously invoke habit.

The three remaining covariates in column 1 capture the city-week-level patterns just discussed. They measure how expensive the expensive fuel is relative to the cheaper substitute—a "magnitude effect"—and how much more expensive the expensive fuel has become relative to the cheaper substitute over the preceding fortnight—an "information diffusion effect."[40] We find that the likelihood that the expensive fuel is chosen over the cheaper one: (i) declines with the magnitude of the price difference, and more so when the cheap fuel is e, and (ii) rises with the recent rate of change of this price difference.

[40] Variable "price premium e above g" is defined as $\max\{p_{el}/k_{ei} - p_{gl}/k_{gi}, 0\}$. The ethanol premium ranges from 0.024 to 0.146 R$/km for the 713 observations where it is strictly positive. Variable "price discount e below g" is defined analogously; the gasoline premium has a lower range, from 0.024 to 0.061 R$/km, for the 289 dear-gasoline observations. Notice that the price premia are considerable. Variable "2-week variation in relative prices" is defined as $|(p_e - L2(p_e))/k_{ei} - (p_g - L2(p_g))/k_{gi}|$, with notation as defined in note 30. We note that relative-price variation for São Paulo averages 0.028 R$/km leading up to January 11 but less than 0.001 R$/km leading up to January 25.

Relative to column 1, column 2 adds city fixed effects. The effect on each of the two price-difference variables is now identified off within-city cross-station price dispersion, and no longer across cities, and as such becomes weaker. The effect on the recent-relative-price-variation variable is identified off only same-city temporal variation and, though significant, is weaker. Other effects are robust. Column 3 replaces city fixed effects and the price difference variables by city-week fixed effects. Again, there is suggestive evidence of some information diffusion: the *p*-value of an equality test between the estimated São Paulo, January 11 fixed effect and its lower São Paulo, January 25 counterpart is 0.06. Finally, columns 4 and 5 indicate that these findings are robust to alternative restrictions that prices be "sufficiently unequal" as well as relative energy differences assumed by consumers.[41]

8. Concluding remarks

This paper has adopted a direct and transparent empirical strategy to uncover substantial consumer heterogeneity in the choice between century-old gasoline and a less-established alternative motor fuel, despite the alternative at hand being: (i) ubiquitously distributed, through the same fueling stations, and over several decades; (ii) purchasable in liquid form at the same point of sale, often at the same pump, comparably priced and billed, linearly per unit volume; and (iii) almost identically consumed, burned in a conventional Otto-cycle combustion engine, operating essentially the same vehicle. One can argue that Brazil's existing FFV-cum-ethanol "chicken *and* egg" substitute is about the closest that can be offered on the supply side to the standard captive gasoline technology. Yet we find demand to segment in observed and unobserved ways. We speculate that the heterogeneous response we have identified is likely to generalize to other markets. Consumer demand for gasoline may prove to be sticky. Substantial investments to educate consumers on alternative fuel technologies may be required. Our study revealed no common perception on the differential effects of gasoline and ethanol with regard to the vehicle's performance and durability, despite the relatively extended period in operation—though it did reveal a significant bias in favor of the "incumbent" gasoline among the subset of motorists who stated that concern with the engine factored into their choice of fuel.

Acknowledgments

This paper previously circulated under the title "Consumer choice between gasoline and sugarcane ethanol." We thank Alcir Bandeira and Denise Guirau from CNP for their professionalism. For facilitating access to data, we are grateful to Claudio Alzuguir (Conpet/Petrobras), Carmen Araujo (Instituto de Energia e Meio Ambiente), Filipe Ferreira (Agência Nacional do Petróleo), and Deise Schiavon (Magneti Marelli & "Circulating Fleet Study Group"). We also thank Soren Anderson, Severin Borenstein, Jim Bushnell, Meghan Busse, Maja Butovich, Chris Conlon, Lucas Davis, David Dranove, Shane Greenstein, Kate Ho, Tom Hubbard, Ben Jones, Kazuya Kawamura, Ryan Kellogg, Maria Kopsakangas-Savolainen, Mike Mazzeo, Mike Riordan, Catherine Thomas, Anna Widerberg, Chris Wilson, Catherine Wolfram, Tim Van Zandt, Florian Zettelmeyer, David Zilberman, as well as audiences at numerous conferences and seminars, for comments. We are also grateful to the editor and to the referees for their thoughtful comments. The project was developed while Salvo was visiting Yale SOM, whom he thanks for their hospitality, thanking Fiona Scott Morton in particular. All errors are our own.

Appendix A. Survey design and data

A.1. Survey details and summary statistics

Market research firm CNP was recommended to us by a seasoned marketing executive in Brazil. Among many consumer choice environments, CNP had experience conducting fieldwork in retail fueling stations. The agreements we signed with CNP, as well as the station-level and motorist-level forms that their field representatives filled out, are available upon request.

Sample of stations. In December 2009 we sent CNP a list of candidate retail fueling stations for each city. For logistical reasons that included the possibility that a station manager's would not authorize a visit and the difficulty in getting around amid traffic congestion, CNP had requested several candidate stations for every station to be visited. We asked that CNP sample at most one station from a given neighborhood (*bairro*), in an attempt to spread stations out in space and raise price dispersion in our sample. We prepared our list of candidate stations from ANP's large retail fueling station database, essentially by restricting stations to be "branded" and to have "purchase invoices" on file. Our intent was to control for branded versus non-branded perceptions of product quality. Most stations in these cities retailed fuel under one among only a handful of brands. For perspective, 72% of stations in the sample we used to prepare Fig. 2 were classified by ANP as "branded," with the brands BR, Shell, Ipiranga, Esso and Chevron accounting for 20%, 15%, 12%, 8% and 7% of the total. ANP's database, and thus our list of candidate stations, did not inform on the availability of midgrade gasoline.

[41] Column 4 uses observations that satisfy $\max\{(p_{gl}/k_{gi})/(p_{el}/k_{ei}),(p_{el}/k_{ei})/(p_{gl}/k_{gi})\}\geq 1.1$, i.e., per-km fuel prices were beyond 10% of each other, such that the number of observations is similar to that in the other regressions. Column 5 restricts observations to those satisfying $|p_{el}/p_{gl}-0.7|\geq 0.06$, i.e., p_{el}/p_{gl} lay at least 6ppt away from the media-reported 70% parity threshold.

The design of a station visit. On arriving at a station, CNP's field representative, wearing a "CNP Research" ("*CNP Pesquisa*") shirt and name tag (ID), would present herself to the station's manager and servicemen (19 out of a total of 21 CNP representatives were women; servicemen, still the norm in Brazil, are typically men). She would emphasize that the "aim of the research was to collect information about FFV motorists, not about the station." On the *station-level form*, she would enter (i) her name and company ID, the date and start time of the station visit, subsequently completing the form with the visit's end time; (ii) the station's name, address, brand, and contact person name and phone number; (iii) per-liter prices of regular gasoline and of ethanol—both fuels g and e were in stock at all visited stations—and of any upmarket gasoline varieties carried by the station at that time (midgrade, $\overline{g}$, premium, $\breve{g}$); and (iv) the number of nozzles dispensing each of these fuel varieties. As the visit progressed, the representative would record a tally of the motorists she approached who appeared to meet certain FFV-qualifying criteria (detailed next) but who declined to participate in the short complementary interview. The proportion of approached consumers who refused participation ranged from 4% in (notoriously friendly northeastern) Recife to 28% in (less friendly southeastern) São Paulo.

FFV-qualifying criteria and sequencing between observations. Each station visit was completed with 12 observations—"FFV-qualifying" motorists—in sequence. A consumer's vehicle had to meet two qualifying criteria: (i) it was originally manufactured as an FFV, i.e., it was neither a single-fuel vehicle, nor was it subsequently retrofitted;[42] and (ii) the FFV was driven for private use, i.e., the car was not a cab or driven at a company's service).[43] Each sampled motorist was asked to confirm both criteria during the interview, though in practice spotting an originally manufactured private-use FFV was not difficult,[44] as newer passenger cars were overwhelmingly FFVs, were typically labeled "FLEX" at the rear or on the side, and did not have company logos or taxi signs attached to them.

Our instruction to the representative was that, on completing observation i, $i<12$, she move on to what seemed to be the next qualifying FFV that pulled up alongside *any* of the station's islands of pumps. Though one same island of pumps typically dispenses both g and e, our concern was to avoid having the representative stand still by an island in which this might not be the case; only observing choices made by FFV motorists who fueled at that particular island would then bias our results. The sequencing rule was "to move to the next FFV motorist by order of arrival *anywhere* in the station." Our data suggests, and casual evidence confirms, that on average no more than one FFV typically fueled at a station at any given moment, providing an incentive for the representative to indeed move around the station (see below).

An observation. Each observation was completed in two stages: (i) passively observing the FFV motorist's choice of fuel, i.e., discreetly watching as the consumer placed his order with the serviceman, and (ii) actively interviewing the motorist while his vehicle was being serviced. We did not want our survey to influence the consumer's choice, observed in the first stage. On approaching the motorist, in the second stage, the representative would follow a pre-set script and introduce herself, express her research aim and state that she would only take a few minutes of the (typically idly sitting) consumer's time to "ask his opinion."

If the motorist was willing to take some short questions, *and* verbally confirmed that his FFV met the two qualifying criteria, the representative would then start making entries on (the next) one of the station's 12 *motorist-level forms*. Having read two filter questions out loud and entered the motorist's qualifying responses (F1. "original flex" → continue, and F2. "private use" → continue), the representative would cover the following questions in sequence, reading the questions out loud (except where noted otherwise) and entering the consumer's responses:

Q1, Q2 (Revealed) Confirmation of the fuel variety and value purchased on that occasion, e.g., ethanol, R$ 50. We instructed the representative to cross-check the consumer's response against the choice that was observable at the pump.

Q3 (Stated) "What was the main reason for your choice of fuel?" ("*Qual o principal motivo de sua escolha do combustível?*") The representative would not read out a menu of possible answers, as the dry runs we conducted had suggested this might frame the motorist's response and take too long. Instead, relying on some cognitive skill, the representative would tick one of five pre-set options that best approximated the consumer's response (e.g., that the consumer was motivated by price), or enter the response verbatim in a sixth category labeled "Other," which we might later reclassify (see below). A seventh category was "I don't know."

Q4 (Stated) "On your last two fueling occasions, what were your fuel choices?" If judged helpful, the representative would provide a hint based on the possible responses, such as "Were your choices both ethanol, both gasoline or one of each?"

Q5 (Stated) "On your last two fueling occasions, how often did you fuel at this station?" The representative would tick one of three possible responses.

Q6 (Stated) "On average, how many kilometers do you drive per week?" The representative would enter the response in a blank cell marked "km/week" or, in view of the question's relative complexity, tick the alternative "I don't know."

[42] Though infrequent, we decided to control for the retrofit possibility by excluding it. Data on the presumably higher-variance fuel economy of such modified engines are not available.

[43] We restricted the population of our study to this majority segment, though understanding the fuel preferences of taxi drivers and firms may in themselves be interesting research questions.

[44] We base this statement also on dry runs we personally conducted in stations in São Paulo and in Rio de Janeiro. We were impressed by the friendliness of local servicemen who would spontaneously call us in their direction as the next originally manufactured private-use FFV pulled up.

Finally, the representative would enter further vehicle and consumer characteristics:

Q7 (Revealed) (i) The vehicle's make and model, including engine size if labeled at the car's rear or on its side: the representative would enter a blank cell, under which we had provided examples, e.g., VW/Polo/1.6; and (ii) the consumer's gender: the representative would tick one of two options;

Q8 (Stated) (iii) The consumer's age and educational attainment: the representative would tick the corresponding category, such as "Over 65 years" and "College, Complete;" and (iv) the consumer's first name and a contact phone number.

"*Internal controls*". According to CNP, requesting consumers' contact details at the close of interviews—characteristic Q8, (iv)—was customary in these local markets. They anticipated from the nature of our survey—learning from consumers rather than attempting to sell them a product—that most consumers would oblige.[45] It was company policy for a "verifier" to contact 20% of the sample to verify that these individuals had indeed participated in the survey, and field representatives were aware of such a policy. It was also common knowledge that during the course of station visits each representative would receive at least one surprise audit from their local CNP supervisor.

Our ideal scenario for the station visit was one in which there was some, but not too much, idle time between motorist observations. Too frequent FFV arrivals might provide an incentive for the representative to rush the data collection, as well as stand by the same island of pumps rather than follow the sequencing rule discussed above. Too sparse FFV arrivals could lead to motivational issues. The mean, minimum and maximum durations of a station visit turned out to be 2.5 h, 50 min and 5.4 h, respectively. Given an expected interview length of under 5 min multiplied by 12 observations per visit, it appears that our ideal scenario largely prevailed.

Descriptive statistics of the sample. Table A1 summarizes station-level data collected in our fieldwork. Variable descriptions are self-explanatory, hence our comments are selective. The first rows describe per-liter regular gasoline and ethanol prices, p_g and p_e, respectively, and the ratio p_e/p_g, at all 180 stations visited. There was considerable cross-sectional and time-series variation in p_e/p_g; this ratio is summarized for each of the nine sampled city-weeks. "Shelf space" allocated to regular varieties g and e hardly changed from January to March, as measured by the mean number of nozzles dispensing each fuel, for the two cities that were surveyed in both months. Compared to the 2.5-hour mean duration of a station visit, a representative's mean time traveling between stations on the same day was (where applicable) a similarly plausible 1.8 h, stuck in traffic or taking time off before returning to work that same day. On average, less than three motorists per station visit refused to take questions (their choices were not recorded).

Table A2 summarizes data that varies across consumers. 45% of the 2160 FFV motorists in our sample ordered (only) g at the pump, whereas 44% chose (only) e. 9% of motorists chose (only) $\overline{g}$, and this share rises slightly to 10% if we condition on the midgrade variety being available at the station. A mere 0.1% of sampled consumers chose $\breve{g}$, corresponding to a still low 2% of choices in the 20 stations where $\breve{g}$ was available. 2% of motorists purchased a "combo" of g and e.[46] Other combos were negligible. The mean fueling ticket was R$ 47: R$ 24 of g, R$ 17 of e, R$ 6 of $\overline{g}$, averaged over all choices. The mean spend among motorists who chose e-only was R$ 37, whereas the mean spend among motorists who chose g-only was (a statistically significantly higher) R$ 52. The observed expenditure distribution translates into a median order of 22 l (not reported in the table). Combined with stated vehicle usage (the mean response among the 1835 motorists who were able to provide an estimate was 296 km/week), the median station stopping frequency is once every 7 days (not reported).

On the dummy variables indicating "the main reason" invoked spontaneously by a consumer for his revealed choice of fuel, three comments are in order. First, the sum of the means for the five precoded-reason dummies, the "Other" dummy and the "I don't know" dummy listed in Table A2 exceeds 1 since almost one-fifth of consumers volunteered more that one reason. Irrespective of the question's wording, we had instructed the representative to record as many reasons as stated by the consumer.

Second, with the wisdom of hindsight, it seems to us that the wording of the second most-marked reason—see the table below—is *not* distinct from the notion of price: while "*tem mais autonomia*" translates into the distinct notion of *Range*, the component "*roda mais*" which loosely translates into "travels further" may have been confused with the top-cited price motivation. Three-quarters of motorists for whom this reason was ticked purchased up to half a tank of fuel, supporting the alternative interpretation that some of these consumers were motivated by price rather than maximizing the distance capacity of their tank (we obtained tank capacities from Carrosnaweb).

Third, we reclassified 74 of the 172 statements originally marked under the "Other" category. Representatives had been instructed to enter the response verbatim in the event that they were unsure as to how to categorize or interpret it. To illustrate, we interpret statements such as "more economical," "I drive more km per liter" and "ethanol is expensive" as being motivated primarily by a *Price characteristic*, rather than non-price characteristics, and reclassified them accordingly.

[45] Only 8% of the 2160 motorists declined to provide a phone number, enabling us to subsequently base the follow-on phone interviews on 1991 consumers. The distribution of missing phone numbers across representatives within a city, as well as the cross-city variation in missing phone numbers, appear plausible, e.g., a lower 2% of (friendly) Recife consumers did not provide a contact number.

[46] Recall that gasoline varieties retailed in Brazil already contained a 20–25% proportion of pure ethanol by volume (this was mandated at 25% between July 2007 and January 2010, and at 20% from February 2010). The average combo consisted of R$ 24 of (blended) g and R$ 32 of e. We instructed representatives to ignore any gasoline used to fill up the 1-liter auxiliary gasoline startup tank that older FFVs came equipped with, which comes into use when starting the engine on ethanol in cold weather, notably around July in the south. Since the value of g recorded on combos was at least R$ 10, or about 4 l, combo observations cannot be attributed to the 1-liter startup tank.

Table A1
Summary statistics at the station-visit level.

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
p_g (R$/l)	180	2.515	0.146	2.199	2.989
January visits only	140	2.550	0.125	2.199	2.989
March visits only	40	2.391	0.148	2.229	2.989
p_e (R$/l)	180	1.879	0.316	1.239	2.549
January visits only	140	2.017	0.196	1.690	2.549
March visits only	40	1.399	0.129	1.239	1.899
p_e/p_g (ratio)	180	0.745	0.105	0.520	1.050
São Paulo on week of January 11 only	20	0.742	0.027	0.692	0.800
São Paulo on week of January 25 only	20	0.747	0.032	0.692	0.833
Curitiba on week of January 25 only	20	0.747	0.021	0.692	0.775
Recife on week of January 25 only	20	0.745	0.022	0.698	0.784
Rio de Janeiro on week of January 25 only	20	0.805	0.037	0.739	0.873
Belo Horizonte on week of January 25 only	20	0.846	0.026	0.800	0.885
Porto Alegre on week of January 25 only	20	0.905	0.041	0.858	1.050
São Paulo on week of March 29 only	20	0.591	0.038	0.520	0.667
Curitiba on week of March 29 only	20	0.578	0.021	0.541	0.609
$p_{\overline{g}}$ (R$/l)	164	2.618	0.185	2.250	3.499
$p_{\overline{g}}/p_g$ (ratio)	164	1.039	0.030	1.000	1.171
$p_{\breve{g}}$ (R$/l)	20	3.213	0.240	2.799	3.699
$p_{\breve{g}}/p_{\overline{g}}$ (ratio)	20	1.162	0.093	1.034	1.375
$p_{\breve{g}}/p_g$ (ratio)	20	1.223	0.095	1.071	1.435
Number of nozzles, all ethanol and gasoline fuels	180	12.522	6.386	3	48
Number of *g*-nozzles	180	5.044	2.422	1	16
São Paulo & Curitiba on week of January 25 only	40	4.875	2.053	2	12
São Paulo & Curitiba on week of March 29 only	40	5.175	2.591	2	16
Number of *e*-nozzles	180	3.900	2.302	1	16
São Paulo & Curitiba on week of January 25 only	40	3.950	2.183	1	10
São Paulo & Curitiba on week of March 29 only	40	4.500	2.792	1	16
Number of $\overline{g}$ nozzles if $\overline{g}$ is available	164	3.677	2.419	1	16
Number of $\breve{g}$ nozzles if $\breve{g}$ is available	20	2.050	0.510	1	4
Fueling station is branded (DV)	180	0.989	0.105	0	1
Fueling station is BR branded (DV)	180	0.294	0.457	0	1
Fueling station is Shell branded (DV)	180	0.267	0.443	0	1
Fueling station is Ipiranga branded (DV)	180	0.189	0.393	0	1
Value of 12 cars sampled in station visit (R$ m)	180	0.403	0.047	0.293	0.571
Station visit began on a weekday peak hour (DV)	180	0.317	0.466	0	1
Station visit began on a weekday non-peak hour (DV)	180	0.467	0.500	0	1
Station visit occurred on a Saturday (DV)	180	0.217	0.413	0	1
Duration of station visit (hours)	180	2.543	0.971	0.833	5.417
Representative's travel time between stations (hours)	100	1.804	1.448	0.133	7.750
Number of consumers who refused to participate	180	2.844	4.026	0	22

Note: Data directly collected in, or constructed from, station-level forms. An observation is a sampled station visit. "DV" denotes dummy variable.

Statements remaining in *Other* in the table below are either orthogonal to the precoded menu of five reasons (such as the 56 consumers who invoked "habit," "accustomed" or a similar reason), or ambiguous or uninformative (e.g., "cost benefit," "family recommendation," and "my option")[47]:

Main reason(s) behind fuel choice as stated by consumer (and subject to representative's and our interpretation)	Number of reasons	
	Raw data	Post reclassification
1 **Price characteristic** "*O preço está mais em conta*"	1454	1475
2 **Range (or Price)** "*Tem mais autonomia/roda mais*"	566	567
3 **Environment** "*Émelhor para o meio-ambiente*"	121	121
4 **Engine performance** "*Quero limpar o motor*"	72	101
5 **Engine startup** "*O motor pega melhor*"	175	179
Other (enter verbatim) "*Outro 1 (descrever)*"	172	98
I don't know "*Não Sei*"	7	7
Total number of reasons:	2567	2548

[47] The total number of reasons drops from 2567 to 2548 since some statements that were entered as "Other" merely reinforce another reason that was already recorded for the same consumer.

Table A2
Summary statistics at the consumer level.

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Purchased g (DV)	2160	0.451	0.498	0	1
January transactions only	1680	0.520	0.500	0	1
Purchased e (DV)	2160	0.435	0.496	0	1
March transactions only	480	0.719	0.450	0	1
Purchased $\overline{g}$ (DV)	2160	0.087	0.282	0	1
Purchased $\breve{g}$ (DV)	2160	0.001	0.037	0	1
Purchased a combo of g and e (DV)	2160	0.024	0.152	0	1
Purchased any other combo (DV)	2160	0.001	0.037	0	1
Value of fuel purchased (R$)	2160	46.973	29.601	10	158
Value of g purchased (R$)	2160	24.052	33.354	0	150
if purchasing only g	975	51.932	31.659	10	150
Value of e purchased (R$)	2160	17.081	23.877	0	140
if purchasing only e	940	37.459	22.401	10	140
Value of $\overline{g}$ purchased (R$)	2160	5.752	21.126	0	150
if purchasing only $\overline{g}$	188	65.477	34.472	10	150
On last two stops, twice chose gasoline (DV)	2160	0.479	0.500	0	1
if now purchasing only g	975	0.825	0.380	0	1
On last two stops, twice chose ethanol (DV)	2160	0.372	0.484	0	1
if now purchasing only e	940	0.778	0.416	0	1
On last two stops, alternated ethanol, gasol. (DV)	2160	0.149	0.356	0	1
Stated reason is Price (DV)	2160	0.683	0.465	0	1
Stated reason is Range or Price (DV)	2160	0.263	0.440	0	1
Stated reason is Environment (DV)	2160	0.056	0.230	0	1
Stated reason is Engine performance (DV)	2160	0.047	0.211	0	1
Stated reason is Engine startup (DV)	2160	0.083	0.276	0	1
Other stated reason (DV)	2160	0.045	0.208	0	1
Does not know the main reason (DV)	2160	0.003	0.057	0	1
Vehicle usage (km/week)	1835	296.094	319.930	5	3500
Of last two stops, twice at this station (DV)	2160	0.513	0.500	0	1
Of last two stops, once at this station (DV)	2160	0.219	0.413	0	1
Of last two stops, none at this station (DV)	2160	0.269	0.444	0	1
Female (DV)	2160	0.342	0.475	0	1
Aged 25–40 years (DV)	2160	0.463	0.499	0	1
Aged 40–65 years (DV)	2160	0.395	0.489	0	1
Aged more than 65 years (DV)	2160	0.037	0.188	0	1
Completed primary school only (DV)	2160	0.059	0.235	0	1
Dropped out of secondary school (DV)	2160	0.026	0.160	0	1
Completed secondary school only (DV)	2160	0.284	0.451	0	1
To complete college education (DV)	2160	0.121	0.326	0	1
Completed college education (DV)	2160	0.497	0.500	0	1
Provided a contact number (DV)	2160	0.922	0.269	0	1
Vehicle price (R$ × 100)	2160	33.602	10.859	21.216	90.732
Vehicle make is Fiat (DV)	2160	0.281	0.449	0	1
Vehicle make is General Motors (DV)	2160	0.248	0.432	0	1
Vehicle make is Volkswagen (DV)	2160	0.231	0.422	0	1
Vehicle make is Ford (DV)	2160	0.102	0.303	0	1

Note: Data directly collected in, or constructed from, motorist-level forms. An observation is a sampled FFV-qualifying motorist purchasing fuel at a station. “DV” denotes dummy variable.

Validating collected data. Vehicle model and engine size required limited cleaning and imputing. The recorded vehicle characteristics reassure us that the qualifying criterion “originally manufactured FFV” was met, since in the raw data there was only one observation for which the recorded model had been discontinued by the automaker prior to FFVs being introduced in 2003 (namely, we changed a Ford Escort record to the newer Ford Ecosport, since Ford discontinued the Escort in 1995). There was also a single observation for which by early 2010 the recorded model had not yet been originally equipped with a flex engine (namely, a Toyota Hilux; one possibility is that the observed vehicle was a retrofit and should have been, but was not, filtered).

We obtained detailed data on the nationwide vehicle fleet circulating at the end of 2009 from the autoparts industry trade association (Sindipeças). This was available by vehicle make, model, year and fuel type but, unfortunately, not at the regional level. We then compared the proportion of models in our six-city survey sample to those in the nationwide stock of FFVs (passenger cars, including SUVs and small/medium pickup trucks). Compared to the nationwide stock, of three leading models, our unweighted survey records oversample the Fiat Palio (13% of our observations vs. 12% of the stock, in all its flex versions), undersample the VW Gol (10% vs. 13%), and sample the GM Corsa at the nationally representative rate (8% vs. 8%). Should data be available on vehicles circulating in each of the sampled cities or neighborhoods by time of day, one could formally test the distribution of vehicle models in our survey records. In the absence of such data, we did not detect any anomaly.[48]

At the more aggregate automaker level, our survey slightly oversamples GM (25% of our observations vs. 23% of the nationwide stock) and Ford (10% vs. 9%), and undersamples Fiat (28% vs. 29%) and VW (23% vs. 27%).

As for the recorded demographic composition of motorists, an old survey by São Paulo city's traffic authority [8] reported that: (i) a majority of the city's motorists were men, though—at 84%—a higher majority compared to our survey decades later; and (ii) 83% of motorists were aged between 25 and 60 years and 4% were over 60 years old, similar to our survey. That half of our sample declared having a college degree, one should note that we were targeting a relatively affluent population of existing consumers, driving a fairly new passenger car on private business (FFVs in circulation were aged on average 2.6 years, according to the Sindipeças data). Also, the definition of college is quite inclusive in that quality varies widely. Our understanding is that consumers were unlikely to misreport their educational attainment category.

We also validated several station-visit level fuel prices in our sample against prices from an external source, ANP's database.

A.2. Other data sources

Vehicle-specific fuel economy. We use laboratory measurements published by Inmetro to predict the fuel economy of the FFVs in our sample. Following US EPA guidelines, Inmetro reports km/l (KPL) values for retailed fuels g (E22) and e (E100) under separate driving cycles—urban and highway—for a fairly broad selection of new FFVs. Note that the cleaning additives in $\overline{g}$ do not materially affect its KPL efficiency relative to g, and that e includes, as retailed, a residual water content, thus known as "hydrated" ethanol. One potential concern is that the cars that were tested in the lab were new, unlike those sampled in the field, some of which had been operating since 2003. However, important for our purpose of analyzing fuel choice, we assume that the KPL *ratio* k_{ei}/k_{gi} does not vary systematically with vehicle use, condition or age.

Specifically, lab measures were available on 23 2009 models and on 44 2010 models. Since measures were not available for several FFVs in our survey sample, we took an indirect route. Using the available lab data, we first projected fuel economy on other vehicle characteristics; we then used the fitted regression to predict each surveyed vehicle's fuel economy based on its observed characteristics. Table A3 presents these auxiliary regressions. We predict six vehicle-specific fuel economy variables, respectively labeled in columns "i" to "vi": k_e^{urban}/k_g^{urban} (i.e., urban driving), $k_e^{highway}/k_g^{highway}$, k_g^{urban}, k_e^{urban}, $k_g^{highway}$ and $k_e^{highway}$. Each dependent variable is projected on two sets of explanatory variables: specification "a" in the upper part of the table or specification "b" in the lower part of the table. When examining consumer choice, we consider urban driving as the baseline and use specification "a," which includes model fixed effects, to predict fuel economy for those surveyed vehicle models which were tested in the lab, i.e., for which there is a model fixed effect such as the GM Celta 1.0 l. This is the case for 66% of FFVs in our survey sample. For the remaining vehicle models that were not tested in the lab, we use specification "b," which includes fixed effects for vehicle make and for vehicle segment. What this means in practice is that we use lab data on similar vehicle models, e.g., the compact VW Fox 1.6 came up in our survey but was not tested in the lab, so we use the estimated VW and compact segment fixed effects.

The rest of the table is self-explanatory, but it is worth pointing out that the R^2 are quite high, in the 75–89% range for specification "a" under urban driving. The fitted value for k_e^{urban}/k_g^{urban}, evaluated at the mean of the regressors in the lab sample, is 0.677 (and 0.681 excluding Renault models, all of which exhibited a KPL ratio below 0.65), with robust standard error of 0.1–0.2%. The KPL ratio is similar across FFVs: the interdecile range (p90–p10) is 0.033, i.e., 0.694–0.661. The mean predicted KPL ratio for surveyed vehicles, adjusting for slightly varying blended gasoline composition between January and March 2010, is 0.687, not shown in Table A3. This indicates that the media-reported parity threshold of 70%, while on the high side, is not wide off the mark.

Vehicle-specific prices. We matched surveyed vehicles to an external dataset on used vehicle prices (Quatro Rodas/Molicar/Fipe). Based on observed make and model (we do not observe the vehicle's age), we took the secondary market price in February 2010 for the most recent model year. For example, Fiat was still producing its Palio 1.0 in 2010 so we took the price of a 2010 Fiat Palio 1.0 in the February 2010 secondary market. On the other hand, GM stopped producing its Meriva 1.8 in 2008 so we took the price of a 2008 GM Meriva 1.8 in February 2010. Since in practice automakers discontinued models slowly, we were able to match nine-tenth of our observations with 2009 or 2010 model prices. The mean, minimum and maximum prices across the 2160 FFVs in our sample are R$ 33,602, R$ 21,216 and R$ 90,732, respectively.

Reassuringly, the mean price of vehicles sampled at stations that carried midgrade gasoline, R$ 31,784, is statistically significantly higher than the mean price of vehicles sampled at the few stations that did not carry midgrade gasoline, R$ 30,834, presumably located in poorer neighborhoods (the p-value of a test of the equality of means is 0.03). In a similar vein, vehicles sampled in relatively less affluent Recife, in the country's Northeast region, have a lower mean price than vehicles

[48] The distribution of vehicle models varies significantly across cities in our survey sample. Unsurprisingly, a χ^2 goodness of fit test indicates that model composition in our unweighted, six-city sample differs significantly from that in the nationwide stock, though differences do not appear to be large.

Table A3
Fuel economy regressions.

Dependent variable	k_e^{urban}/k_g^{urban}	k_e^{highw}/k_g^{highw}	k_e^{urban}	k_e^{urban}	k_e^{highw}	k_e^{highw}
Specification	[a-i] coeff. (s.e.)	[a-ii] coeff. (s.e.)	[a-iii] coeff. (s.e.)	[a-iv] coeff. (s.e.)	[a-v] coeff. (s.e.)	[a-vi] coeff. (s.e.)
Size of the engine (l)	−0.011**	−0.004	−1.036***	−0.814***	−0.576*	−0.430**
	(0.005)	(0.009)	(0.233)	(0.158)	(0.312)	(0.184)
Transmission is non-manual (DV)	−0.013	−0.007	−0.015	−0.147	0.251	0.072
	(0.008)	(0.009)	(0.167)	(0.104)	(0.216)	(0.142)
Intercept	0.709***	0.684***	11.250***	7.940***	12.156***	8.302***
	(0.009)	(0.013)	(0.388)	(0.282)	(0.438)	(0.257)
Vehicle model fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.750	0.629	0.883	0.893	0.914	0.938
Fitted value at covariate mean (in lab data)	0.677	0.674	10.393	7.036	12.652	8.528
	(0.002)	(0.002)	(0.043)	(0.028)	(0.057)	(0.033)
Specification	[b-i] coeff. (s.e.)	[b-ii] coeff. (s.e.)	[b-iii] coeff. (s.e.)	[b-iv] coeff. (s.e.)	[b-v] coeff. (s.e.)	[b-vi] coeff. (s.e.)
Size of the engine (l)	−0.012**	−0.003	−1.530***	−1.170***	−0.956**	−0.685**
	(0.005)	(0.008)	(0.330)	(0.239)	(0.463)	(0.337)
Transmission is non-manual (DV)	−0.013**	−0.006	0.013	−0.122	0.145	0.016
	(0.006)	(0.007)	(0.147)	(0.098)	(0.222)	(0.148)
Vehicle segment is compact (DV)	−0.007*	−0.006	−0.553*	−0.454**	−0.536	−0.431
	(0.004)	(0.006)	(0.295)	(0.208)	(0.380)	(0.262)
Vehicle segment is midsize (DV)	−0.007	−0.004	−0.263	−0.257	−0.447	−0.350
	(0.004)	(0.007)	(0.295)	(0.216)	(0.421)	(0.293)
Vehicle segment is full-size (DV)	0.002	−0.003	−0.925**	−0.594**	0.138	0.064
	(0.015)	(0.015)	(0.371)	(0.253)	(0.534)	(0.372)
Vehicle segment is small truck (DV)	−0.011**	−0.004	−0.659	−0.571*	−1.656**	−1.174**
	(0.005)	(0.008)	(0.455)	(0.307)	(0.737)	(0.542)
Intercept	0.710***	0.689***	12.685***	8.945***	13.636***	9.369***
	(0.008)	(0.011)	(0.580)	(0.416)	(0.759)	(0.529)
Vehicle make fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.674	0.468	0.553	0.562	0.556	0.519
Fitted value at covariate mean (in lab data)	0.677	0.674	10.393	7.036	12.652	8.528
	(0.001)	(0.002)	(0.070)	(0.048)	(0.110)	(0.078)
Number of observations	67	67	67	67	67	67

Note: Estimated coefficients from auxiliary OLS regressions that predict vehicle-specific fuel economy based on laboratory test data reported by Inmetro. An observation is a vehicle model in the Inmetro test sample. In "a" specifications, vehicle model fixed effects are included (the Fiat Palio subcompact dummy variable is omitted). In "b" specifications, vehicle make fixed effects and vehicle segment fixed effects are included (the Fiat and the subcompact dummy variables are omitted). "DV" denotes dummy variable. Heteroscedasticity-robust standard errors in parentheses.

* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

sampled in the five southern/southeastern cities, with mean prices of R$ 30,834 and R$ 33,948, respectively (an equality test yields a p-value of 0.00).

Since zip-code-level incomes are not available, we take the combined price of the 12 vehicles sampled during each station visit as a proxy for the average income of a station's customers. In addition to cross-city variation, there is substantial within-city variation, e.g., R$ 0.29–0.44 million in less affluent Recife, R$ 0.34–0.57 million in more affluent Rio de Janeiro: some neighborhoods, or motorists along certain routes, are more affluent than others. To support the notion that variation in station-level vehicle prices signals spatial differences in income distribution and not simply variation in sampling from a common income distribution, we can exploit the fact that 34 stations in São Paulo and Curitiba were each visited multiple times. The aggregate price of the 12 vehicles surveyed during each visit for a station visited multiple times varies little compared to cross-station variation.

Appendix B. Phone-based interview evidence

The table below details the responses of $N' = 607$ motorists, during follow-on phone interviews by a professional at market research firm InHouse, to the following three questions that attempted to gauge their views of gasoline against ethanol:

Q1′ "In your view, the consumption of which motor fuel pollutes less and is better for the environment: gasoline, ethanol or is there no difference?" ("*Na sua opinião, o consumo de qual combustível polui menos e é melhor para o meio ambiente, gasolina ou álcool ou não faz diferença?*"). Recall that each interview was randomly assigned an order in which the two fuels would be stated in the questions, i.e., either gasoline followed by ethanol or ethanol followed by gasoline.

Q2′ "In your view, the consumption of which motor fuel is better for the performance of your Flex car: gasoline, ethanol or is there no difference?" ("*Na sua opinião, o consumo de qual combustível é melhor para o desempenho de seu carro Flex, gasolina ou álcool ou não faz diferença?*"). Consumers responding gasoline or ethanol were immediately asked whether their FFV had experienced problems on burning ethanol or gasoline, respectively. In both cases, more than 75% of respondents said they had not experienced problems—we omit these responses for brevity.

Q3′ "In your view, the consumption of which motor fuel is better for Brazil: gasoline, ethanol or is there no difference?" ("*Na sua opinião, o consumo de qual combustível é melhor para o Brasil, gasolina ou álcool ou não faz diferença?*"). Motorists responding gasoline or ethanol were immediately asked "Why?" ("*Por quê?*"), with no assistance from the interviewer—we briefly report answers in Section 5.

Favorable views comparative	Motorists observed in station	Total phone respond	"Environ." Q1′→ethanol[a]		"Vehicle performance"				"Country" Q3′→ethanol[a]	
					Q2′→ethanol[a]		Q2′→gasoline[b]			
São Paulo	720	190	156	82%	52	27%	100	53%	142	75%
Curitiba	480	151	129	85%	61	40%	57	38%	112	74%
Recife	240	75	66	88%	23	31%	40	53%	52	69%
Subtotal *e* producers	1440	416	351	84%	136	33%	197	47%	306	74%
Rio de Janeiro	240	76	64	84%	20	26%	40	53%	49	64%
Belo Horizonte	240	66	50	76%	21	32%	36	55%	38	58%
Porto Alegre	240	49	35	71%	10	20%	27	55%	29	59%
Subtotal *e* importers	720	191	149	78%	51	27%	103	54%	116	61%
Total	2160	607	500	82%	187	31%	300	49%	422	70%

[a] Number and proportion of total phone respondents who selected ethanol.
[b] Number and proportion of total phone respondents who selected gasoline.

The following table evaluates the answers of 403 phone respondents who, as mentioned in Section 5, were asked "How do you calculate whether ethanol is favorably or unfavorably priced?" ("*Como o/a senhor/a calcula se o Álcool está caro ou barato?*").

Understanding & application of price parity	Number respond	% of 403	Cum. %
Good or reasonable description of (or reference to) 70% parity ratio	136	34	34
Good or reasonable description of price heuristic for ethanol $0.7p_{gl}$	28	7	41
Reasonably alludes to "cost per km" even if briefly	54	13	54
Refers to "use of table" (internet/online calculator/smartphone app, etc.)	25	6	60
Appears confused with respect to, or does not remember, cost conversion[a]	39	10	70
Motorist states or suggests that he/she does not do, or know, cost conversion[a]	86	21	91
Avoids directly answering question	35	9	100

[a] Fuel choice might still be guided by, e.g., the radio, colleagues or the serviceman.

Appendix C. Aggregate fuel share predictions

We use the estimated individual-level choice model to predict aggregate fuel shares for the three richest Brazilian states, comparing these predictions against available market-level reports of fuel quantities. In principle, our survey design accounts for varying rates of vehicle usage in the population, since extensive drivers are more likely to be sampled relative to less extensive users. In practice, since single-fuel vehicle users were not surveyed, one needs additional information to account for this subpopulation, such as from a household-level travel study, e.g., the NHTS in the US. In the absence of such data for Brazil, we make some assumptions, as our purpose here is to illustrate.

The southeastern states of São Paulo, Minas Gerais and Rio de Janeiro—respectively SP, MG and RJ—together account for about 40% of the country's population and over 50% of its GDP. Since the metropolitan areas of Brazilian state capitals are home to a disproportionate share of state populations, we assume that a state capital's motorists (and cars), whom we sampled, are fairly representative of motorists in the wider state, e.g., São Paulo metro's population accounts for about half of SP state residents. Following Salvo and Huse [23], we base the composition of a state's circulating passenger-vehicle fleet on data on new vehicle registrations (available by state but not by engine type), new vehicle sales (available by engine type) and scrappage rates. For example, SP state's fleet in February 2010 by engine type was 56% gasoline-captive (mostly pre-2005), 11% ethanol-captive (pre-2005) and 33% flex-fuel (introduced in 2003).

The more "heroic" assumption in this illustration regards relative vehicle usage across engine types (vintage): here we merely assume that the ratio of vehicle kilometers traveled (VKT) to km per liter (KPL) does not vary significantly between single-fuel vehicle users and FFVs operating on the same fuel. One rationale could be that relative to their single-fuel counterparts, more-modern dual-fuel vehicles are more efficient but at the same time are utilized more as their owners tend to be wealthier. For perspective, absent data, MMA [18] assumes that annual VKT is 20,000 km for new cars and 2000 km for 30-year-old (and substantially less efficient) cars still in circulation.

To predict fuel-choice shares among FFV motorists, we feed the weekly evolution of state-level pump prices between December 2009 and May 2010 into specification I (Table 2). We hold consumer characteristics constant at the empirical distributions observed in the state capital's survey subsample, and weigh choice probabilities across consumers using the distance embedded in their observed purchases. The table below reports mean per-liter price ratios and market-level shares, both for the subset of FFV motorists and for the population of car drivers, on selected weeks. For example, in the first week of March 2010, facing p_e/p_g of 83%, 20% of Minas Gerais' 1 m FFVs refueled with ethanol.

Selected week (2009/2010)	December 7 (%)	January 18 (%)	March 1 (%)	April 12 (%)	May 24 (%)
(Per-liter) p_e/p_g, SP	64	74	71	63	55
(Per-liter) p_e/p_g, MG	73	80	83	73	72
(Per-liter) p_e/p_g, RJ	71	79	80	70	69
Predicted ethanol share of gasoline-plus-ethanol "vehicle km purchased:"					
FFVs only, SP (2.9 m cars)	62	49	50	61	73
FFVs only, MG (1.0 m cars)	33	25	20	31	34
FFVs only, RJ (0.6 m cars)	42	30	29	42	45
All passenger cars[a], SP (8.9 m cars)	31	27	28	32	36
All passenger cars[a], MG (2.6 m cars)	22	20	18	22	23
All passenger cars[a], RJ (2.1 m cars)	24	21	20	24	26

Note: February 2010 fleet estimates.
[a] Assumes that VKT:KPL is equal across flex- and single-fuel.

How do these predictions compare against any available market-level shipment data? ANP compiles monthly distributor reports of shipments to retailers, allegedly by state of destination. We caution that, as a measure of local fuel consumption, such data might not be comprehensive, or fully capture interstate shipments (e.g., out of ethanol-producing SP state), or account for variation in downstream inventories. We find that temporal variation in the reported ethanol share of gasoline-plus-ethanol shipments, in energy-adjusted "barrels of oil equivalent," is somewhat consistent with the last three rows of the table (for all cars): ethanol's reported shipment share across the three states was 38% in December 2009, bottoming out at 25% in February 2010, and rising to 38% by May 2010. That these reported shares tend to exceed our predicted entire-fleet shares may also reflect a bias in the "heroic" relative-usage assumption we used to form an aggregate prediction: in contrast to the assumption, perhaps newer FFVs burn more liters than single-fuel cars, i.e., the average VKT:KPL ratio for a typical flex-fuel car operation may be higher than for a single-fuel one. Predicted shares align more closely with reported ones if we use the more-price-sensitive specification III (Table 2) estimates, without city fixed effects.

References

[1] Agência Estado, Indústria inicia campanha para incentivar uso do etanol, Edition of November 8 (Industry Launches Campaign to Promote the Use of Ethanol), 2012 (in Portuguese).
[2] S.T. Anderson, The demand for ethanol as a gasoline substitute, Journal of Environmental Economics and Management 62 (2012) 151–168.
[3] S.T. Anderson, J.M. Sallee, Using loopholes to reveal the marginal cost of regulation: the case of fuel-economy standards, American Economic Review 101 (2011) 1375–1409.
[4] I. Ayres, S. Raseman, A. Shih, Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage, Journal of Law, Economics, and Organization, forthcoming, 〈http://jleo.oxfordjournals.org/content/early/2012/08/18/jleo.ews020.abstract?papetoc〉, http://dx.doi.org/10.1093/jleo/ews020.
[5] B.J. Bronnenberg, J.-P.H. Dubé, M. Gentzkow, The evolution of brand preferences: evidence from consumer migration, American Economic Review 102 (2012) 2472–2508.
[6] R. Casadesus-Masanell, M. Crooke, F. Reinhardt, V. Vasisth, Households' willingness to pay for "green" goods: evidence from Patagonia's introduction of organic cotton sportswear, Journal of Economics & Management Strategy 18 (2009) 203–233.
[7] S. Clerides, P. Courty, Sales, Quantity Surcharge, and Consumer Inattention, CEPR Discussion Paper 8115, 2010.
[8] CET, Companhia de Engenharia de Tráfego, Caracterização do motorista paulistano segundo sexo, faixa etária e tempo de habilitação, Notas Técnicas 121/121, Traffic Authority of the city of São Paulo, Brazil (Characteristics of São Paulo City's Motorists: Gender, Age Category and Driving Experience), 1987 (in Portuguese).
[9] K.S. Corts, Building out alternative fuel retail infrastructure: government fleet spillovers in E85, Journal of Environmental Economics and Management 59 (2010) 219–234.
[10] Economist, The Fat of the Land: Green-Minded Motorists are Making Car Fuel at Home, from Used Cooking Oil, Edition of December 3, 2011.
[11] D.W. Elfenbein, B. McManus, A greater price for a greater good? Evidence that consumers pay more for charity-linked products, American Economic Journal: Economic Policy 2 (2010) 28–60.
[12] A. Goolsbee, A. Petrin, The consumer gains from direct broadcast satellites and the competition with cable TV, Econometrica 72 (2004) 351–381.
[13] Grist, 15 Green Cities, Posted online on July 19 (Accessed at 〈http://www.grist.org/article/cities3〉), 2007.
[14] S.P. Holland, J.E. Hughes, C.R. Knittel, Greenhouse gas reductions under low carbon fuel standards? American Economic Journal: Economic Policy 1 (2009) 106–146.
[15] J.E. Hughes, C.R. Knittel, D. Sperling, Evidence of a shift in the short-run price elasticity of gasoline demand, Energy Journal 29 (2008) 113–134.
[16] M.S. Jacobson, Effects of ethanol (E85) versus gasoline on cancer and mortality in the United States, Environmental Science & Technology 41 (2007) 4150–4157.

[17] Richard G. Lugar, US Senator and Chairman of the Senate Foreign Relations Committee, Open letter to the editor of the Wall Street Journal, Edition of February 13, 2006.
[18] MMA, Ministério do Meio Ambiente, Primeiro inventário nacional de emissões atmosféricas por veículos automotores rodoviários, Ministry for the Environment, Brasília, Brazil (First National Inventory of Atmospheric Emissions Released by Automobiles), 2011 (in Portuguese).
[19] E.J. Miravete, Choosing the wrong calling plan? Ignorance and learning, American Economic Review 93 (2003) 297–310.
[20] New York Times, US Apple Growers Feel Heat from China, Edition of June 25, 2007.
[21] Quatro Rodas, Exclusivo! Testamos o primeiro carro a álcool e gasolina. Edition of April 2003, (Exclusive! We Tested the First Car that Burns Ethanol and Gasoline), 2003 (in Portuguese).
[22] E.M. Rogers, Diffusion of Innovations, fourth ed. The Free Press, New York, 1995.
[23] A. Salvo, C. Huse, Is arbitrage tying the price of ethanol to that of gasoline? Evidence from the uptake of flexible-fuel technology, Energy Journal 32 (2011) 119–148.
[24] United States Environmental Protection Agency, EPA Lifecycle Analysis of Greenhouse Gas Emissions from Renewable Fuels, Regulatory Announcement EPA-420-F-10-006, February 2010.
[25] Wall Street Journal, India's Sugar Crop Wilts Amid Drought, Edition of September 9, 2009.