

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

Alberto Salvo

Northwestern University (Kellogg)

Cristian Huse

Stockholm School of Economics

May 2012*

Abstract

How motorists 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 among century-old gasoline and a less-established—but still widely available and usable—alternative, 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.

JEL classification: D12, D64, L62, L71, Q21, Q41, Q42, R41

Keywords: Fossil fuels, renewable fuels, alternative fuels, gasoline, ethanol, biofuels, flexible-fuel vehicles, consumer choice, consumer preferences, perfect substitutes, differentiated products

*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. The project was developed while Salvo was visiting Yale SOM, whom he thanks for their hospitality, thanking Fiona Scott Morton in particular. The usual disclaimer applies.

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 fuel 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 fairly “similar” ones 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 motorists 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. 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., Wall Street Journal 2009).

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

¹For example, with regard to ethanol in the U.S., Corts (2010) writes that “(w)hat the required discount to gas(oline) is remains a debated topic.” (p.7)

²This is a common modelling assumption, e.g., Holland et al (2009), Salvo and Huse (2011).

per-liter price of gasoline p_g . That is, prices equalize when $p_e/p_g \simeq 70\%$, 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 $p_e/p_g = 70\%/1.2 \simeq 58\%$. Similarly, we observe roughly 20% of FFV motorists choosing ethanol when ethanol is priced 20% above gasoline in \$/km, i.e., when $p_e/p_g = 70\% \times 1.2 = 84\%$.

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 85% (with this ratio varying little across stations), translating to 0.241 Brazilian Real (R\$) per km traveled on gasoline against a 22%-dearer 0.294 R\$/km on ethanol.⁴ Despite the hefty ethanol price premium, we observed 49 motorists, or 20.4% of the sample, choosing ethanol over gasoline. Importantly, ethanol had been more expensive than gasoline for *over six weeks*, and in our data the median motorist refuels once a week. We argue that motorists had ample additional opportunities to become informed about effective fuel prices via, for example: (i) the local radio, with relative fuel price newscasts losing only to traffic updates in terms of frequency, (ii) the station serviceman, who fuels the vehicle and can offer price advice upon request (one can view him as a Bayesian updater), 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 motor fuel. To illustrate why this finding is important for energy planning, we conduct a counterfactual, predicting aggregate fuel shares were state sales tax on ethanol in the remote Amazonian state of Pará⁵ 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 Bronnenberg et al 2012). But the 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.

³Many motorists 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 $0.7p_g$.

⁴To follow the fuel economy math, notice that $0.294/0.241 = 1.22 \simeq 85\%/70\%$.

⁵In 2010, Pará’s urban population was 5m and FFVs accounted for 45% of its light vehicle fleet.

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 (likely) wealthier motorists 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 motorists 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, motorists voicing environmental concerns or residents in sugarcane-growing states are more likely to choose the locally produced renewable fuel.

The closest paper to ours is Anderson (2012), 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 (favorably priced ethanol is not observed, unlike in our setting), and the sparser availability of FFVs and ethanol poses selection challenges. Anderson and Sallee (2011) 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 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.⁶ 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 predominantly drives 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 (2009) and Elfenbein and McManus (2010) (also see references therein) examine whether consumers will actually pay more for substitute products perceived to

⁶For example, conditional on buying any Volkswagen car model as of 2006, a motorist would acquire an FFV (Salvo and Huse 2011). 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 100-200 US\$ (Corts 2010, Anderson and Sallee 2011), possibly not worth the cost of carrying different engines.

be associated with “good causes” or linked to charity. We find strong evidence of this: some motorists 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.⁷ A related literature examines whether households are sensitive to their neighbors’ energy conservation efforts (e.g., Ayres et al 2011). 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 (Rogers 1995).⁸

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 summary statistics, including empirical demand curves. Section 4 analyzes consumer response by way of multinomial probit models. Section 5 briefly presents evidence from follow-on phone interviews. Section 6 provides a counterfactual to illustrate why our demand results are relevant to energy supply planners. Section 6 concludes by further considering the relevance of our findings.

2 Conjectured departures from perfect substitution

We briefly conjecture why substitution between gasoline and ethanol—by FFV motorists who at least in principle can seamlessly substitute—might be less than “perfect.” To fix ideas, let k_f denote a vehicle’s fuel economy in kilometers per liter of either fuel $f \in \{g, e\}$, and recall that p_f denotes the consumer price in \$ per liter. Consider a motorist whose vehicle usage is M km per period, a year say. (For brevity, in this section we include subscripts i to index motorists only on a selective basis.) Annualized expenditure on fuel f is then Mp_f/k_f . Motorists choosing only based on fuel prices would choose f such that $p_f/k_f = \min\{p_g/k_g, p_e/k_e\}$. We refer to such behavior, for labeling purposes only, as “perfect substitution.”

Begin with the time cost of stopping to refuel. Suppose that this cost is linear in the number of stops and that a motorist incurs S \$ per stop. Denoting the vehicle’s tank capacity in liters by T , and assuming that the motorist depletes and fills up the tank every time (e.g., that the motorist does not face liquidity or cash-in-pocket constraints), the annualized stopping cost of refueling on fuel f is $MS/(Tk_f)$. This stopping cost makes motorists, all else equal, favor gasoline over ethanol, since $k_g > k_e$. In particular,

⁷The 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)” (US EPA 2010; parentheses added for clarity). In contrast, when it comes to local pollutants—i.e., urban air quality—rather than global emissions, Jacobson (2007) finds that the use of ethanol is not superior to gasoline.

⁸Subsequent 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 (2003) and Clerides and Courty (2010).

we conjecture that wealthier motorists might exhibit a higher stopping cost S_i and thus be more likely to choose gasoline over ethanol relative to less affluent motorists.

Alternatively, suppose that some motorists exhibit convex, rather than linear, stopping costs. For example, consumers might be “loyal” to their local fueling station⁹ (e.g., through repeated interaction they trust that their neighborhood station does not dilute the fuels it sells with solvents or water) or have a favored fueling occasion (e.g., on the way back from, but not to, work or school), making them averse to fueling outside their favored station or occasion. Motorists who regularly commute on highways may be particularly prone to “range anxiety.” We conjecture that such convex stopping costs make heavy commuters (those with higher M_i) more likely to choose gasoline over ethanol relative to motorists who use their vehicles less extensively, since, all else equal, the former are more likely to find themselves stopping to refuel outside their favored shopping location or occasion, thus valuing the extra range afforded by a tank of gasoline (i.e., $T(k_g - k_e)$).

Conjecture 1 [*Time Costs and Wealth*] *Time stopping costs may make wealthier motorists more likely to choose gasoline over ethanol relative to less affluent motorists.*

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 light commuters.*

Now consider differential maintenance costs as perceived by motorists. Despite ethanol fuel having been distributed and used widely for a quarter century, some motorists might associate the more established “secular” gasoline fuel with lower lifetime maintenance costs or improved performance. We subsequently report evidence of this. Again to fix ideas, denote by C_f^{mainten} the perceived (and additive) maintenance cost per km traveled on fuel f ; annualized maintenance costs would then be MC_f^{mainten} . To the extent that some motorists perceive $C_{g,i}^{\text{mainten}} < C_{e,i}^{\text{mainten}}$ they would, all else equal, favor gasoline over ethanol.

Further, following the technology adoption literature (Rogers 1995), 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 vehicles in the 1980s. For example, the tanks of early ethanol vehicles were not lined with alcohol-resistant polymers so back then these were more prone to corrosion compared to gasoline vehicles.

Conjecture 3 [*Technological Concerns*] *Motorists voicing technological concerns as the primary reason for their choice of fuel may be more likely to choose gasoline over ethanol relative to other motorists.*

⁹We find evidence of this, as we subsequently report.

Conjecture 4 [*Age and Technology Adoption*] *Older motorists may be more likely to choose gasoline over ethanol relative to younger motorists.*

Next consider perceived effects on the natural environment. As we subsequently report, 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 motorists respond ethanol. Some of these consumers may intrinsically care enough for the environment and decide to take matters into their own hands.¹⁰ Denote by C_f^{environ} the perceived *and* internalized cost to the environment from traveling 1 km on fuel f . We conjecture that consumers who are more susceptible to environmental concerns—i.e., for whom $C_{g,i}^{\text{environ}} > C_{e,i}^{\text{environ}}$ —are more likely to choose the renewable fuel ethanol over the fossil fuel gasoline.

Finally, consider perceived differential effects on the local economic environment. In a similar vein to the natural environment, motorists 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 motorists perceive greater benefits to the home economy from consuming ethanol *and* they display home bias in their fuel choices—i.e., they internalize “per-km benefits” $B_{e,i}^{\text{home}} > B_{g,i}^{\text{home}}$ —they may, all else equal, favor “locally-sourced” ethanol over “imported” gasoline.¹¹ 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 5 [*Environmental Concerns*] *Motorists voicing environmental concerns as the primary reason for their choice of fuel may be more likely to choose ethanol over gasoline relative to other motorists.*

Conjecture 6 [*Home Bias*] *Motorists who reside in sugarcane-growing states may be more likely to choose ethanol over gasoline relative to motorists who reside in ethanol-importing states.*

To sum up, the annualized cost for motorist i from choosing fuel f (in the linear

¹⁰For example, “the Baltimore Biodiesel Co-op, in Maryland, 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” (Economist 2011). In our setting, driving off with ethanol hidden away in the tank is not an act of conspicuous consumption.

¹¹This follows decades of pro-ethanol advertising by the sugar industry and the government. (In view of Brazil’s recent deepwater oil discoveries, such views may change over the coming decades.) Says U.S. 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 U.S. all the time” (New York Times 2007).

stopping cost case) is given by:

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

While not meant to provide a theory of consumer choice between alternative energy sources, this expression illustrates some “taste-based” tradeoffs that (price-attentive) motorists may make on choosing between gasoline and ethanol. It is also plausible that heavy commuters, who make extensive use of their vehicles, or less affluent motorists, who on average drive less expensive vehicles, 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. Figures 1 and 2 depict our empirical strategy.

Figure 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 (lower panel). All prices are in constant Brazilian Reals (R\$). Two points are important. First, local fossil fuel prices are still 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 at 150 US\$/bbl in mid-2008, were not passed through to the gas pump. Second, and in contrast to the first point, market forces have been at work in Brazil’s sugar/ethanol industry: the opportunity cost of selling ethanol (or sugar) on domestic markets is given by the export price of sugar.¹² During this decade, whenever world sugar prices crossed a certain threshold—say, 0.40 R\$/lb, and this did happen in early 2003, 2006 and 2010—ethanol prices in São Paulo fueling stations reached 2 R\$/liter. (The exchange rate stood at 1.9 R\$/US\$ in early 2010, so divide by two for rough prices in US\$.)

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—namely to observe and interview FFV motorists making choices at fueling stations—when so instructed by us, at informative price points in time. We adopted a consumer-level approach rather than estimate demand from market-level data, on two counts. First, our aim was to examine heterogeneity among subsets of the population. Second, while good-quality local price data were available, market-level quantity data were either less reliable or

¹²See Salvo and Huse (2011) 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.

unavailable.¹³

Figure 2 hones in on the 2009/10 relative fuel price fluctuation in a cross-section of six major cities where we ultimately observed consumer choices. The four curves in each panel indicate different percentiles (p5, p25, p75, p95) of the distribution of the ethanol-to-gasoline price ratio, p_e/p_g , in a large weekly sample of a city’s retail fueling stations, according to a publicly-available source, the National Oil Agency (ANP). To be clear, these are prices per liter for the regular variety of each fuel at the *same* station, and often times at the same pump (though different nozzle), given the ubiquitous supply of both gasoline and ethanol. The nine vertical lines mark the nine “city-weeks” 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).

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 and ethanol producer, Porto Alegre is 2000 km from the nearest sugarcane plantation.¹⁴ Notice also that *within-city* relative price dispersion was low compared with relative price variation across cities—consider the interquartile range for p_e/p_g , marked by the two inner curves in each panel of Figure 2.

Perfect substitution and the media What does this variation mean in terms of energy-adjusted prices? To illustrate, consider the fuel economy 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—necessarily “Flex”—operated under an “urban driving cycle” produces 9.9 km/liter of gasoline and, given ethanol’s lower energy content, 6.9 km/liter of ethanol. For this particular car (and fuel composition), $k_{e,i}/k_{g,i} = 6.9/9.9 \simeq 70\%$.

This 70% relative “price parity” ratio is indeed regularly reported in the media, particularly at times when ethanol prices vary, as in 2009 Qtr 4 and 2010 Qtr 1, not least because ANP prices for many cities are so readily accessible online. Intense media

¹³For example, FFV fleet size and usage relative to older single-fuel cars, by state or city.

¹⁴Salvo and Huse (2011) show that all three states where São Paulo, Curitiba and Recife are located feature among only 8 states of Brazil (out of a total of 27) whose share of the national sugarcane harvest exceeded their share of national GDP. For example, in 2005-07 São Paulo accounted for 61% of the country’s harvested sugarcane and 34% of its GDP, suggesting that the state was a “net producer” of ethanol (ethanol mills locate close to plantations). By contrast, all three of the other surveyed states were “net consumers” (or importers) of ethanol. Salvo and Huse also report that, 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 support their local sugar industry (or penalize it less) by way of a lower state sales tax on ethanol—see counterfactuals.

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 stands relative to 70%.

A hypothetical Fiat Palio user who chooses fuel on the basis of prices alone would then have purchased gasoline in the January 2010 city-weeks of our sample—since $p_e/p_g > k_{e,i}/k_{g,i}$ or, equivalently, $p_g/k_{g,i} < p_e/k_{e,i}$. By March 2010, in São Paulo or Curitiba, this motorist would have refueled with ethanol. Even in Porto Alegre, ethanol had been favorably priced relative to gasoline as recently as September 2009. As we explain below, the median $k_{e,i}/k_{g,i}$ across FFVs in our field sample is 69%, and there is little variation around this ratio.¹⁵

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 were a fortnight apart: the rise in p_e/p_g petered out at 75%, prior to dropping. Our aim was to use these observations in an attempt to gauge short-run habit persistence, which we conjecture to be an additional reason—on top of the “taste-based” ones enumerated above—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 wished to collect 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 $9 \times 20 \times 12 = 2160$ consumer-level observations, with each city-week totaling 240 observations. As we detail in Appendix C, 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 (we naturally left the market research firm in charge of the specific visit logistics).

By design, we required that regular gasoline g and ethanol e (*gasolina comum* and *álcool comum*, respectively) be available during each station visit, a constraint that in practice tends to be non-binding. Most stations also retailed “midgrade” gasoline (*gasolina aditivada*), which we denote by \bar{g} , and a few stations sold the further upmarket “premium” gasoline variety (*gasolina premium*), which we denote by \check{g} .

We instructed the field representative conducting a visit to discreetly observe each FFV motorist’s fuel choice among the alternatives available at the station, and only then—once the station’s serviceman had begun servicing the vehicle, inside which the motorist was typically sitting idly—approach the motorist for a short interview with the

¹⁵We detail the source of model-specific fuel economy data in Appendix C. For gasoline, Inmetro publishes $k_{g,i}$ for a blend containing a 22% ethanol component, i.e., E22. Since gasoline as retailed in 2010 was actually an E25 blend in January and an E20 blend in March (having changed by federal mandate), we linearly (and slightly) adjust $k_{g,i}$. Retailed ethanol was unblended E100 all along.

aim to collect information about the motorist and the vehicle. The questions included confirming that the car was an FFV and driven for private use (“filters” which in practice are easy to spot); asking about “the main reason” for the motorist’s choice of fuel observed on that occasion (without showing a menu of options, which might frame the response); and inquiring about the motorist’s car usage and age and schooling categories; 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, until a total of 12 observations were completed per station visit, as mentioned.

Summary statistics We keep this subsection brief and refer the reader to Appendix C for further details. Figure 3 indicates the location of retail fueling stations visited by our field representatives in each of the six cities in January and/or March. The mean number of standard-variety nozzles across the $9 \times 20 = 180$ station visits was five g -nozzles and four e -nozzles; also, “shelf space” of g and e did not vary between January and March visits. Midgrade gasoline \bar{g} was available on 91% of station visits, retailing at a mean (same-station) markup of 4% over g . Premium gasoline \check{g} was available on only 11% of station visits, at a mean 16% markup over \bar{g} .

In our sample ($N = 2160$), the “median motorist” is male (66% of the sample), states being middle-aged (25-40y 46% and 40-65y 40%), states having attained higher education (50% completed a college degree), spontaneously invokes “price” as the primary motivation for his fuel choice (68%), and stops to refuel once a week. (We estimate the stopping frequency from the observed fueling amount, the stated car usage and the model-specific fuel economy under urban driving.) Fifty-one percent (resp., 73%) of the sample stated that of the last three fuel purchases, including the present one, they had fueled at that particular station thrice (resp., at least twice), suggesting that in equilibrium “shopping around” does not substantially occur in these markets.

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 (discussed next), 83% stated having purchased gasoline on both of their two immediately preceding station stops (with the caveat that this particular statistic is stated). Similarly, among motorists whom we observed purchasing ethanol, 78% stated having chosen ethanol on both of their two preceding stops. Returning to Figure 2, note that by the week of January 25 (six city-weeks in our sample) the ethanol price hike had mostly already occurred. Similarly, by the week of March 29 (two city-weeks), p_e/p_g had already dipped below the (approximate) 70% price parity threshold two to three weeks

earlier.¹⁶

Empirical demand Figure 4 summarizes the choices made by FFV motorists by aggregating these to the station level. January station visits are marked with circles and March observations are marked with squares. We plot the station’s per-liter ethanol price relative to regular gasoline, p_e/p_g , against ethanol’s overall “share” in the 12 choices we observed in that visit, computing this share in two ways. In the left panel, for each station visit we count the number of FFV motorists who chose ethanol over gasoline as their dominant source of energy (“fuel km purchased”) on that fueling occasion, and divide by 12; that is, we define the ethanol share as:

$$s_{e,j} = \frac{1}{12} \sum_{i \in \mathcal{O}_j} \mathbf{1} \left[q_{e,i} \frac{\widehat{k}_e}{\widehat{k}_{g_i}} > \sum_{f \in \{g, \bar{g}, \check{g}\}} q_{f,i} \right], \quad (2)$$

where \mathcal{O}_j is the set of motorists observed during station visit j , $\mathbf{1}[x]$ is an indicator function (equal to 1 if condition x holds and 0 otherwise), and the observed quantity in liters of ethanol purchased by motorist i , $q_{e,i}$, is adjusted for ethanol’s lower kilometerage per liter (kpl) relative to gasoline liters purchased. Appendix C explains how we predict the kpl ratio, under urban driving, for every vehicle in our field sample: the median $\widehat{k}_e/\widehat{k}_{g_i}$ is 69.0%, and the median standard error is 0.5% (that is, the prediction is fairly tight, precise within about 1.0%). Clearly, this “unweighted” ethanol share is discrete-valued, as it is based on a count variable.¹⁷

The right panel of Figure 4 alternatively computes ethanol’s share of the aggregate kilometerage purchased by the 12 FFV motorists observed in the station visit:

$$\tilde{s}_{e,j} = \left(\sum_{i \in \mathcal{O}_j} q_{ei} \frac{\widehat{k}_e}{\widehat{k}_{g_i}} \right) / \sum_{i \in \mathcal{O}_j} \left(q_{ei} \frac{\widehat{k}_e}{\widehat{k}_{g_i}} + \sum_{f \in \{g, \bar{g}, \check{g}\}} q_{fi} \right). \quad (3)$$

Like its unweighted counterpart, this “weighted” ethanol share indicates that there is considerable consumer heterogeneity. Departure from perfect substitution is significant.

A close look at either panel of Figure 4 reveals that facing relative prices p_e/p_g of around $70/1.2 \approx 58\%$ (i.e., g priced roughly 20% above e per km traveled), about one-fifth of motorists stay with gasoline. Similarly, facing p_e/p_g of $70 \times 1.2 \approx 84\%$ (e priced 20% above g), about one-fifth of motorists choose ethanol.

¹⁶Returning to the two São Paulo-January subsamples that are a fortnight apart but exhibit the same relative prices (p_e/p_g rising, then flattening out), it is noteworthy that among FFV motorists observed purchasing gasoline the week of January 11, 69% stated having fueled with gasoline on their two preceding stops, to be compared with a higher 85% of motorists observed purchasing gasoline the week of January 25 who stated having fueled with gasoline on their two preceding stops.

¹⁷We do not employ the simpler condition $q_{e,i} > 0$ in (2) since 2.5% of motorists 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 simpler condition would, however, yield a similar plot.

One may ask how much of this consumer heterogeneity is explained by variation in the kpl ratio $k_{e,i}/k_{g,i}$ across FFVs (noting that the interdecile range in the surveyed sample is 3.8%). By controlling for “parity” differences across car models, Figure 5 shows that the answer is “not much.” To plot the figure, we compute the difference $p_{e,i}/p_{g,i} - \widehat{k_e/k_{g_i}}$ for each of the 2160 observations; we then collect observations in one percentage point bins and compute the share of motorists who chose ethanol rather than gasoline as their dominant energy source (defined per the condition in the indicator function of (2)). To illustrate by way of a data point, the motorist of a VW Gol 1.0 Flex (another popular car) fueling at a particular Belo Horizonte station in January faced a $p_{e,i}/p_{g,i}$ ratio of 88.2% at the pump and a predicted kpl ratio of 69.9% (with s.e. 0.5%); this motorist’s choice would enter the $88.2\% - 69.9\% \simeq 18\%$ bin. Figure 5 indicates that the empirical probability that this motorist would have chosen ethanol over gasoline is still no less than a sizable 10-15%, despite facing distance-adjusted fuel prices p_f/k_f of 0.28 R\$/km on ethanol against a substantially cheaper 0.22 R\$/km on regular gasoline.¹⁸ Computed at the median car usage rate (200 km/week), this 0.06 R\$/km or 27% price premium represents 12 R\$/week, equivalent to an annualized R\$ 624 (i.e., $M(p_e/k_e - p_g/k_g)$) expenditure difference.

Importantly, the heterogeneous consumer response depicted in Figures 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 (see Appendix C), 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 Analyzing consumer choice

We begin by modeling motorist i as having utility

$$u_{f,i} = x_i' \delta_f - \alpha p_{f,i}/k_{f,i} + \varepsilon_{f,i} \quad (4)$$

from choosing fuel $f \in \{g, e, \bar{g}\}$, where $p_{f,i}/k_{f,i}$ is the vehicle-specific price in R\$ per km traveled that the motorist faces at the station for the fuel, vector x_i contains observed motorist and vehicle characteristics, and unobserved tastes $\varepsilon := (\varepsilon_g, \varepsilon_e, \varepsilon_{\bar{g}})$ follow a multivariate Normal distribution with mean zero and covariance matrix Ω , i.e., $\varepsilon \sim MVN(0, \Omega)$. To illustrate, the probability that a motorist fuels with ethanol is

$$\begin{aligned} \Pr(i \text{ chooses } e) &= \Pr(u_{g,i} - u_{e,i} \leq 0 \cap u_{\bar{g},i} - u_{e,i} \leq 0) \\ &= \Phi(-((x_i' \delta_f - \alpha p_{f,i}/k_{f,i}) - (x_i' \delta_e - \alpha p_{e,i}/k_{e,i})), \Omega_{-e}), \quad f = g, \bar{g}, \end{aligned}$$

¹⁸Had we instead plotted $p_{e,i}/\widehat{k_{e,i}} - p_{g,i}/\widehat{k_{g,i}}$ on the vertical axis of Figure 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.

where Φ is the CDF of the bivariate normal random variable $(\varepsilon_g - \varepsilon_e, \varepsilon_{\bar{g}} - \varepsilon_e)$ with mean zero vector and covariance matrix Ω_{-e} .^{19,20} Energy-adjusted prices in our baseline multinomial probit specifications consider predicted fuel economy $k_{f,i}$ under urban driving but estimates are robust to alternative assumptions (more below). We rely on the moderate within-city-week (within-route) price dispersion, coupled with motorists’ professed “station loyalty,” and ignore any substitution across stations. To be clear, what we rule out is, e.g., an observed ethanol consumer substituting e at this station for g at another (unobserved) station with a lower p_g (and thus the relevant p_e/p_g would be higher), an event that would likely reinforce our results.²¹

Table 1 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,²² whereas increasing age is associated with the choice of gasoline. The latter correlation is captured by Conjecture 4 (Age and Technology Adoption) above. In particular, all else equal, motorists aged 65y+ are 27 percentage points (ppt, per column I) less likely to choose ethanol over gasoline compared to motorists aged 25y–, who appear more comfortable with the alternative fuel. Heavy commuters—defined here as motorists whose stated car usage places them in the upper quartile of the empirical car 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 car models—defined as those in the upper quartile of the survey’s distribution of model prices²³—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

¹⁹Thanks 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 discrete choice settings in transportation and industrial organization (e.g., Goolsbee and Petrin 2004 examine competition between cable and satellite TV using data on 30,000 households in 317 markets).

²⁰We place no additional restrictions on Ω beyond those necessary to identify the model, i.e., with $F = 3$ alternatives, Ω has $F(F - 1)/2 - 1 = 2$ free terms, say one error variance parameter $\sigma_{\bar{g}}^2$ and one correlation parameter $\rho_{g,\bar{g}} = \text{Corr}(\varepsilon_{i,g}, \varepsilon_{i,\bar{g}})$. Further, in the 9% of station visits in which midgrade gasoline \bar{g} was not available, motorists are modeled as choosing between two alternatives, g or e .

²¹We further: (i) ignore the availability of premium gasoline \check{g} (11% of station visits), as we observed a mere 3 motorists choosing \check{g} , and reclassify these observations as midgrade gasoline \bar{g} choices; and (ii) reclassify 54 “combo” observations (essentially motorists who ordered both g and e on the same fueling occasion, as per note 17) as choosing a single fuel according to the dominant kilometerage source in the order. Results are robust to dropping, rather than reclassifying, these 57 observations (2.6% of sample).

²²That said, we find that women are three percentage points less likely to choose midgrade gasoline over alternatives than men (not shown; p-value 0.014; \bar{g} ’s mean choice probability in the sample is 7%).

²³See Appendix C on how we matched observed vehicles to estimates of their value. We also discuss why the aggregate price of the 12 vehicles sampled during each station visit, which varies between 0.29 and 0.57 million reais, may reasonably proxy for the average income of a station’s customers.

vehicle models: since expensive (larger, powerful) cars tend to burn more fuel and pollute more than cheaper vehicles,²⁴ “green-minded” types may favor smaller cars and locally grown sugarcane ethanol.²⁵ We attempt to control for neighborhood-level incomes by including the aggregate price of the 12 cars sampled in the motorist’s visited station, but this is not significant.

What does vary across the three specifications of Table 1 is the inclusion of: (i) city fixed effects, in columns I and II, relative to III; and (ii) dummy variables indicating the main reason invoked by a motorist as the basis for his fuel choice, in column II, relative to I and III.

In the presence of city fixed effects (columns 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 price dispersion observed across visited stations (in all cities). With city-level shocks being soaked up by city fixed effects, and our proxy for the average income of a station’s shoppers 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 $p_{f,i}/k_{f,i} - p_{f',i}/k_{f',i}$, 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. Rather than nuisance parameters, we interpret the estimated city fixed effects as indicative of home bias. (Alternatively, one can interpret these estimates as evidence of long-run habit formation.) We find that every one of the three capitals of sugarcane-growing states—São Paulo, Curitiba and Recife—displays a stronger taste for ethanol relative to each of the other ethanol-importing cities—Rio de Janeiro, Belo Horizonte and Porto Alegre—and this difference is significant at the 5% level in *all* nine pairwise tests (per specification I; test statistics are not reported). This result is consistent with Conjecture 6 (Home Bias).

Comparing column II to column I, price and other effects are robust to including stated-reason dummies. One may be willing to take the view that a motorist’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 motorists who, when interviewed shortly after revealing their preference, spontaneously invoked “the environment” or (maintenance or performance of) “the engine.” We do this by interacting stated-reason dummies with relative price levels, namely dummies for an

²⁴In the data, vehicle price correlates tightly (and inversely) with fuel economy, with correlation coefficients of -0.64 between vehicle price and k_g , or -0.63 between price and k_e . That is, pricey cars tend to consume more energy and emit more carbon per km traveled.

²⁵Plausibly, owners of expensive vehicles may also be more inclined to adopt an upmarket variety (available in 91% of visited stations). Indeed we find that such motorists are 2ppt more likely to choose \bar{g} , but the effect is not significant (p-value 0.133).

ethanol price premium of at least 10%, (i.e., $(p_{e,i}/k_{e,i}) / (p_{g,i}/k_{g,i}) \geq 1.1$), a gasoline premium of at least 10%, and an ethanol or gasoline premium under 10%. Our rationale for specifying such interactions is that an environmentally concerned motorist, 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 greenery against price and most other motorists would also choose ethanol—recall (1)).²⁶

On controlling for stated reasons, we find that invoking greenery is associated with a substantial increase in the probability that the motorist’s choice was ethanol over gasoline (unconditionally, in the data, 83% of motorists stating greenery chose 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, motorists 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., Quatro Rodas 2003).

Finally, we investigate how estimates change on dropping city fixed effects—which we interpreted as evidence of home bias (Conjecture 6, Home Bias). Column III shows that 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 switch-

²⁶Results 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 ($N = 2160$) invoke the environment, a lower 4.7% and 4.5% of motorists invoke 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 motorists who invoke the environment on facing moderate (mostly ethanol) premia ($N = 804$).

ing 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 2 shows that the above results are largely robust to plausible variations in relative energy differences assumed by motorists, as well as to using different subsets of the data. Rather than consider “true” vehicle-specific fuel prices $p_{f,i}/k_{f,i}$ in R\$ per km traveled (recall (1) and (4)), 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 $(p_{g,i}, p_{e,i}, p_{\bar{g},i})$, a motorist would denominate these in the same “currency” $(0.7p_{g,i}, p_{e,i}, 0.7p_{\bar{g},i})$ irrespective of his FFV. For example, a motorist facing $p_{g,i}$ and $p_{e,i}$ of 2.729 and 2.199 R\$/liter, respectively, would compare $2.729 \times .7 \simeq 1.910$ to 2.199, applying the heuristic $p_{e,i} < 1.91 \Rightarrow$ “ e priced below g ,” rather than compare $2.729/k_{g,i}$ to $2.199/k_{e,i}$. 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.²⁷ Plausibly, this slightly larger price sensitivity suggests that the widely reported 70% parity ratio may be more ingrained than the particular kpl ratio for each FFV—we subsequently offer phone-based interview evidence of this. Other effects are similar.

Another specification with which we experimented, though omit from Table 2 for lack of space, 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 observed across a city over the previous week (like the radio, we use median $(0.7p_g, p_e)$ for the previous week’s ANP sample). We obtain similar results, indicating that the moderate relative price dispersion across stations within a city—switched off in this unreported specification—is not significantly driving price sensitivity in our baseline results.²⁸

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 in January and March (and prices are again $p_{f,i}/k_{f,i}$). Price effects are estimated off of time variation. To account for the early January reading in São Paulo, immediately following the ethanol price hike, we control for the increase in the (median citywide) ethanol price premium over the preceding two weeks, defining the (weakly positive) covariate $\max\{0, (p_e - L2(p_e))/k_{e,i} - (p_g - L2(p_g))/k_{g,i}\}$, with $L2$ denoting the two-week lag op-

²⁷Sample means for $(0.7p_{g,i}, p_{e,i}, 0.7p_{\bar{g},i})$ are (1.88, 1.76, 1.83) R\$/liter, not shown in Table 2 due to space constraints.

²⁸Price effects under this “newscast” specification are slightly higher than in column IV, offering suggestive evidence that some motorists complement prices they read at the pump with more salient information reported on the local radio. This robustness test—and several others that are not reported (e.g., controlling for a fuel’s “shelf space” at the station)—are available upon request.

erator.²⁹ For perspective, this “recent rate of increase in the relative ethanol price” averages 0.027 R\$/km for the São Paulo January 11 subsample 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 motorists—it is not significant, neither statistically nor economically. Indeed, in the raw data, the proportion of São Paulo motorists paying 8% more per km for ethanol was a not-too-dissimilar 51% in early January compared to 42% a fortnight later. Importantly, effects (price and non-price) 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 two 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, or 15% of the sample, stated they did not know their usage). This specification can be viewed as very flexibly introducing random coefficients on price and non-price characteristics in (4), based on stated vehicle usage. The price sensitivity for motorists who profess to use their vehicle more extensively 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 (at or above the median of R\$ 29,504 versus below the median), as a proxy for motorist’s socioeconomic standing—see the two columns marked “VIIa” and “VIIb.” Price and non-price effects are similar to specification I; if anything, motorists of cheaper vehicles appear slightly more price sensitive. Results suggest that our baseline specification is not excessively restrictive.³⁰

Illustrating consumer heterogeneity What choices does an “average” motorist make at different relative prices? Consider a male motorist aged 25-40y, who states having at least some college education, who neither states commuting extensively nor drives an expensive car. In the left panel of Figure 6, we plot choice probabilities for this median motorist using specification III (Table 1) 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

²⁹Appendix B further explores an “information diffusion effect.”

³⁰Other 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\$ 0.37m and R\$ 0.44m per station, respectively. Estimated effects are similar.

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, $p_{e,i}/k_{e,i}$, while holding constant: (i) the per-km regular gasoline price, $p_{g,i}/k_{g,i}$, at 0.246 R\$/km (this is the mean across the entire sample); (ii) its midgrade gasoline counterpart at 0.256 R\$/km; and (iii) the three-alternative choice set $\{g, e, \bar{g}\}$.

As the left panel indicates, when e is priced at parity to g , i.e., $p_{e,i}/k_{e,i} = p_{g,i}/k_{g,i} = 0.246$ R\$/km, the probability of choosing e is just over 60%, and the choice probabilities for g and \bar{g} are just under 35% and 5%, respectively. What is striking is that even for this median motorist, when e is priced at a substantial premium relative to g —say $p_{e,i}/k_{e,i} = 0.316$ R\$/km, equivalent to a $0.070/0.246 = 29\%$ premium over g (i.e., $p_{e,i}/p_{g,i} \simeq 90\%$)—the choice probability for e is still a sizable 19%! Similarly, when e is priced at a substantial discount relative to g —say a 0.070 R\$/km, or 29%, discount ($p_{e,i}/p_{g,i} \simeq 50\%$)—the probability that this motorist still chooses gasoline (regular or midgrade) is a non-negligible 9%.³¹ This heterogeneous response to prices among motorists with a given set of observed characteristics is evidence of considerable unobserved consumer heterogeneity. To draw a contrast with survey data, the right panel of Figure 6 plots choice probabilities estimated off a sample of 2160 motorists facing actual prices and choice sets but whom we *simulate* to care mostly about prices. That is, for this simulation, in utility function (4) we *set*: (i) non-price coefficients on observables at zero, $\delta_f = 0$; (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 Figure 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 25y– with some college education and who resides in Curitiba, the capital of sugarcane-growing, ethanol-producing Paraná state (other motorist/vehicle characteristics are switched off); and (ii) a “gasoline fan,” defined as an older male aged 65y+ with no more than primary education, who commutes extensively, drives an expensive vehicle 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 Figure 7 plots ethanol choice probabilities for consumers with above-median versus below-median stated vehicle usage, as in columns “VIa” and “VIb” of Table 2 (the plot considers otherwise median motorists in a “median” city such as Rio de Janeiro). Relative to less extensive users, motorists who state using their vehicles more extensively are somewhat more price sensitive and, faced with price parity, tend to favor gasoline over ethanol.

³¹Had we based Figure 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 Figure 8 below.

“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 motorist spontaneously invoking the environment to have ordered ethanol (resp., gasoline)? Figure 8 plots choice probabilities for median motorists (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 environment-invoking main-reason dummy from “off” to “on,” marked by the thin and thick lines respectively.³² To illustrate, an environment-invoking Rio motorist facing an ethanol price of 0.34 R\$/km has the same 50% probability of adopting ethanol as a non-environment-invoking Rio motorist 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 ($N' = 607$)

As part of a follow-on project, in late 2010 we hired another market research firm—one that specializes in telephone marketing—to contact motorists in our sample for additional questions; 1991 of the $N = 2160$ motorists observed at the station had provided contact details at the end of their interview (see Appendix C). A phone interviewer worked on the 1991 contacts for over a month and managed to conclude $N' = 607$ phone-based interviews.

After reminding the motorist about the face-to-face interview earlier that year, the very 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 motorists, selected ethanol, and a further 13% felt there was no differential impact on the environment (recall the motivation behind Conjecture 5, Environmental Concerns).

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 422 out of the 607 motorists (i.e., 70%)

³²We illustrate by turning on the moderate price-difference dummy “Envir. & price premium < 10%” 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.

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 motorists who stated that Brazil benefited from substituting ethanol for gasoline were then simply asked “Why?”: 194 motorists (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 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.³³ See Appendix D for all the by-city proportions across the three questions.³⁴

The follow-on phone interviews also probed into motorists’ understanding of price parity across ethanol and gasoline, whether they applied the parity concept by calculating $p_{f,i}/k_{f,i}$ in R\$/km, compared the per-liter $p_{e,i}/p_{g,i}$ ratio against the media-reported 70% rate, or formed the gasoline-based price heuristic for ethanol $0.7p_{g,i}$. All 607 phone participants were asked “How do you decide on which fuel to purchase?”; the 403 motorists whose response was clearly based on price³⁵ 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 D for details.)

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

³³Recall Figure 7. Anecdotally, Curitiba has been labeled Brazil’s “green capital,” e.g., U.S. magazine Grist (2007) places Curitiba third in a list of “15 Green Cities,” after Reykjavik, Iceland and Portland, Oregon, U.S. 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 0.013).

³⁴Two 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 motorist’s phone responses increases the power of the baseline demand specifications (Table 1) in predicting his fuel choice months earlier, e.g., phone respondents who stated that their vehicle performed better on ethanol were 14 to 17ppt (depending on the specification) more likely to have purchased ethanol at the station (1% significance level).

³⁵We did not enquire about parity if a motorist, e.g., voiced a strong preference for one fuel, such as the 63 motorists who stated (something resembling) “I always choose gasoline/I dislike ethanol.”

6 A counterfactual: Planning the energy mix

The estimated discrete-choice model can be used to predict market-level shares for the substitute fuels under different relative price scenarios. Appendix A provides a detailed example for the three richest Brazilian states, comparing predictions—at current prices and under certain assumptions—against available reports of market-level fuel quantities. In what follows, we briefly illustrate what our finding of imperfect demand substitutability might mean for energy planners.

Consider a planner in the Amazonian state of Pará (PA), in the remote north of Brazil, home to 7.6m people in 2010, two-thirds of whom lived in urban areas. Thanks to a high state sales tax on ethanol, coupled with the federal government’s favorable gasoline pricing policy toward remote states,³⁶ PA 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 the approximate 70% parity threshold was in 2002, and since then through May 2010 it hovered between 75% and 90%, with a mean of 82%.³⁷ 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 PA state planner is considering a reduction in the state sales tax (known as ICMS) on ethanol, to shift the source of energy powering the state’s FFVs to ethanol, away from gasoline, say to meet some carbon emissions target.³⁸ 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 motorists were to display the tastes and behavior of no-more-remote Porto Alegre, where historically ethanol has been no more expensive than in Pará. Our demand estimates suggest—see the column marked “Counterfactual 1” in the table below—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%. (We keep the tax calculation simple—quite unlike the reality of Brazilian taxes—as our purpose is to illustrate. We also ignore a supply response from a lower tax—e.g., ethanol producers raising prices—and use the distribution of motorist characteristics and fuel availability in our survey, with the Porto

³⁶Recall note 14. The state is as distant from sugarcane plantations as it is from oil refineries.

³⁷That 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 A) suggest that even this pricey location has its share of “ethanol fans.” Ethanol’s share of Pará’s gasoline-plus-ethanol “fuel km purchased”—across both FFVs and single-fuel (essentially gasoline-dedicated) vehicles—was 4.0% in December 2009 ($p_e/p_g = 78\%$), falling to 3.1% in March 2010 ($p_e/p_g = 83\%$), and rising to 4.3% in May 2010 ($p_e/p_g = 77\%$). We estimate the penetration of ethanol-dedicated vehicles at only 1% of the state’s fleet, thus the bulk of these ethanol shares, as well as their variation, are likely due to FFVs.

³⁸We are not advocating any particular policy, just illustrating the planner’s problem.

Alegre fixed effect turned on.)

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

Notes: 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%. We discuss possible policy prescriptions in our concluding remarks.

7 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 one can get to on the product 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—and perhaps even in a magnified way. For example, in the absence of price discounts, large-scale substitution to electric cars is likely to be slow, given the quite different and varied technologies, the limited range and the less familiar pricing

structure relative to gasoline.³⁹

The results we offer are important, since policies to raise consumption of alternative fuels have largely focused on the supply side, tending to take the demand side for granted. For example, U.S. Senator Richard G. Lugar (2006) argued that “(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.” In addition to other costs, the senator likely did not consider the non-negligible price discount that might be required, or welfare loss that some consumers might experience, were U.S. light transportation to “switch to an ethanol-based system.” Consumer demand for gasoline may prove to be sticky.

Substantial investments to educate consumers on alternative fuel technologies may also be required. Our study revealed no common perception on the differential effects of gasoline and ethanol with regard to the vehicle’s engine power, startup, durability and maintenance costs, despite the relatively extended period in operation (though there is a bias in favor of gasoline among the subset of motorists—12% of the sample—who expressed concern with the engine).

References

- [1] Anderson, S. T. (2012). The demand for ethanol as a gasoline substitute. *Journal of Environmental Economics and Management*, forthcoming
- [2] Anderson, S. T. and J. M. Sallee (2011). Using loopholes to reveal the marginal cost of regulation: The case of fuel-economy standards. *American Economic Review* 101, 1375-1409
- [3] Ayres, I., S. Raseman and A. Shih (2011). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. Yale Law & Economics Research Paper No. 437
- [4] Bronnenberg, B. J., J.-P. H. Dubé and M. Gentzkow (2012). The evolution of brand preferences: Evidence from consumer migration. *American Economic Review*, forthcoming
- [5] Casadesus-Masanell, R., M. Crooke, F. Reinhardt and V. Vasisth (2009). Households’ willingness to pay for “green” goods: Evidence from Patagonia’s introduction of organic cotton sportswear. *Journal of Economics & Management Strategy* 18, 203-233
- [6] Clerides, S. and P. Courty (2010). Sales, quantity surcharge, and consumer inattention. Mimeo, University of Cyprus and University of Victoria

³⁹Further, vehicles are big ticket items, which may slow down trial, particularly in light of possible early supply glitches that may reverberate in the minds of some consumers more than others (e.g., see Miravete and Moral 2011 on Spain’s experience with substitution from gasoline to not-that-dissimilar diesel).

- [7] CET, Companhia de Engenharia de Tráfego (1987). 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, in Portuguese.)
- [8] Corts, K. S. (2010). Building out alternative fuel retail infrastructure: Government fleet spillovers in E85. *Journal of Environmental Economics and Management* 59, 219-234
- [9] Economist (2011). The fat of the land: Green-minded motorists are making car fuel at home, from used cooking oil. Edition of December 3
- [10] Elfenbein, D. W. and B. McManus (2010). A greater price for a greater good? Evidence that consumers pay more for charity-linked products. *American Economic Journal: Economic Policy* 2, 28-60
- [11] Goolsbee, A. and A. Petrin (2004). The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV. *Econometrica* 72, 351-381
- [12] Grist (2007). 15 green cities. Posted online on July 19. (Accessed at <http://www.grist.org/article/cities3>.)
- [13] Holland, S. P., J. E. Hughes and C. R. Knittel (2009). Greenhouse gas reductions under low carbon fuel standards? *American Economic Journal: Economic Policy* 1, 106-146
- [14] Jacobson, M. S. (2007). Effects of ethanol (E85) versus gasoline on cancer and mortality in the United States. *Environmental Science & Technology* 41, 4150-4157
- [15] Lugar, U.S. Senator and Chairman of the Senate Foreign Relations Committee, Richard G. (2006). Open letter to the editor of the Wall Street Journal. Edition of February 13
- [16] MMA, Ministério do Meio Ambiente (2011). 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, in Portuguese.)
- [17] Miravete, E. J. (2003). Choosing the wrong calling plan? Ignorance and learning. *American Economic Review* 93, 297-310
- [18] Miravete, E. J. and M. J. Moral (2011). Qualitative effects of “Cash-for-Clunkers” programs. Mimeo
- [19] New York Times (2007). U.S. apple growers feel heat from China. Edition of June 25
- [20] Quatro Rodas (2003). Exclusivo! Testamos o primeiro carro a álcool e gasolina. Edition of April 2003. (Exclusive! We tested the first car that burns ethanol and gasoline, in Portuguese).
- [21] Rogers, E. M. (1995). *Diffusion of Innovations*. Fourth edition. New York: The Free Press

- [22] Salvo, A. and C. Huse (2011). Is arbitrage tying the price of ethanol to that of gasoline? Evidence from the uptake of flexible-fuel technology. *Energy Journal* 32
- [23] United States Environmental Protection Agency (2010). EPA lifecycle analysis of greenhouse gas emissions from renewable fuels. Regulatory Announcement EPA-420-F-10-006, February
- [24] Wall Street Journal (2009). India’s sugar crop wilts amid drought. Edition of September 9

Appendix: For online publication only

A Appendix: Predicting market-level fuel shares

How can one go about predicting fuel shares at the market level? In principle, our survey design accounts for varying rates of vehicle usage in the population, since extensive commuters are more likely to be sampled relative to less extensive users. In practice, we focused our survey dollars on the (fast-rising) one-third of motorists driving FFVs, so one needs additional information, such as from a household-level travel study (e.g., the NHTS in the US), to account for single-fuel car users. In the absence of such information for Brazil, we make some assumptions, as our purpose here is to illustrate.

In what follows, we consider the three richest Brazilian states—São Paulo (SP), Minas Gerais (MG) and Rio de Janeiro (RJ), all in the Southeast region—which together account for about 40% of the country’s population and over 50% of its GDP. We assume that a state capital’s motorists (and cars), whom we sampled, are fairly representative of motorists in the wider state (the metropolitan areas of Brazilian state capitals are home to a disproportionate share of state populations, e.g., São Paulo metro’s population accounts for about half of SP state residents). 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 car sales (available by engine type) and scrappage rates (see Salvo and Huse 2010 for details). For example, SP state’s car 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 car usage across engine types (vintage): here we merely assume that the ratio of vehicle kilometers traveled (vkt) to kilometerage per liter (kpl) does not vary significantly between single-fuel car users and FFVs operating on the same fuel. One rationale could be that relative to their single-fuel counterparts, more-modern dual-fuel cars have better fuel economy but are utilized more by (on average) wealthier owners.⁴⁰

We feed the weekly evolution of (mean) state-level pump prices between December 2009 and May 2010 into the estimated multinomial probit specification I, holding constant the distribution of motorist characteristics in the state capital’s survey subsample, with the aim to predict fuel-choice shares among FFV motorists on a weekly basis. (Choice probabilities are weighted across motorists by the kilometerage embedded in their observed purchases, in the spirit of (3).) 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 commuters, on selected weeks. For example, in the first week of March 2010, facing p_e/p_g of 83%, 20% of Minas Gerais’ (1m) FFVs refueled with ethanol.

⁴⁰MMA (2011) assumes that annual vkt is 20,000 km for new cars and 2,000 km for 30-year-old cars still in circulation. Besides usage, one would need (sparsely available) fuel economy data for early vintage cars or require a household-level travel study to collect fuel *liters* purchased in addition to vkt.

Selected week (2009/10):	Dec 7	Jan 18	Mar 1	Apr 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 “fuel km purchased:”					
FFVs only, SP (2.9m cars)	62%	49%	50%	61%	73%
FFVs only, MG (1.0m cars)	33%	25%	20%	31%	34%
FFVs only, RJ (0.6m cars)	42%	30%	29%	42%	45%
All passenger cars*, SP (8.9m cars)	31%	27%	28%	32%	36%
All passenger cars*, MG (2.6m cars)	22%	20%	18%	22%	23%
All passenger cars*, RJ (2.1m cars)	24%	21%	20%	24%	26%

Notes: Feb. 2010 fleet estimates. *Assumes vkt:kpl is equal across flex- and single-fuel

How do these predictions compare against any available market-level shipment data? The National Oil Agency (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 Agency’s reported ethanol share of gasoline-plus-ethanol shipments, in energy-adjusted “barrels of oil equivalent” (boe), 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 2010.⁴¹ That these reported shares tend to exceed our predicted entire-fleet shares may also reflect a bias in our “heroic” relative car usage assumption above: unlike what we assumed in these predictions, 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).

Appendix: For online publication only

B Appendix: Choosing the expensive fuel

We further examine heterogeneity by considering the choices of certain motorists in the sample, facing sufficiently unequal prices across the two standard fuels g and e . In such markets, which observable characteristics help explain the choice of the dear fuel over a cheaper substitute?

We restrict the present analysis to motorists who appear less likely to have a strong taste for one fuel. Specifically, we consider motorists who invoked neither environmental nor engine concerns as the basis for their observed fuel choice, *and* who stayed clear of non-standard fuel \bar{g} (and \check{g}). We interpret any motorist who spontaneously expressed concern for a non-price characteristic, or who revealed a preference for \bar{g} over (cheaper) g , as indeed valuing a characteristic other than price, thus dropping these observations from the analysis that follows. (Alternatively, we could have kept these observations and controlled them away.) We look for evidence of “consumer inattention” (or “short-run habit formation”) where we are more likely to find it.

⁴¹Predicted shares align more closely with reported ones if we use the more-price-sensitive specification III estimates, without city fixed effects (as illustrated in Figure 6).

We also discard observations in which motorists faced (broadly) similarly-priced g and e (in R\$/km), since not much can be inferred about consumers' understanding of relative prices from choices made in such markets.

The table below describes the resulting restricted set of 1002 observations and, in the final column, the proportion of these 1002 motorists who chose the dear fuel.

City-week	Number of observations:				Proportion choosing the expensive fuel
	Motorists observed	(-) Facing similar prices*	(-) Likely strong taste	(=) Included in analysis	
São Paulo, Jan 11	240	175	14	51	26/51 = 51%
São Paulo, Jan 25	240	151	17	72	27/72 = 38%
Curitiba, Jan 25	240	145	8	87	44/87 = 51%
Recife, Jan 25	240	174	15	51	33/51 = 65%
Rio de Janeiro, Jan 25	240	32	74	134	44/134 = 33%
Belo Horizonte, Jan 25	240	0	46	194	41/194 = 21%
Porto Alegre, Jan 25	240	0	116	124	10/124 = 8%
São Paulo, Mar 29	240	81	20	139	36/139 = 26%
Curitiba, Mar 29	240	57	33	150	19/150 = 13%
Total	2160	815	343	1002	280/1002 = 28%

Notes: *Observations facing “similar” prices, as reflected in the table and dropped in specifications I to III of Table 3, are those for which $\left| p_{g,i}/\widehat{k}_{g,i} - p_{e,i}/\widehat{k}_{e,i} \right| < 0.0242$ R\$/km, i.e., for which fuel prices were within 0.0242 R\$/km of each other. 0.0242 R\$/km is calculated as one-tenth the mean price of the cheaper fuel, $\min_{f \in \{g,e\}} \left\{ p_{g,i}/\widehat{k}_{g,i}, p_{e,i}/\widehat{k}_{e,i} \right\}$, where the mean is taken over all 2160 motorists i in the survey sample.

Before presenting probit estimates of the choice between the dear fuel (g or e) and the cheap alternative, we comment on some intuitive city-week-level data patterns that will inform our specifications.

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 (a mean p_e/p_g of 74.2% in the early subsample against 74.7% two weeks later), yet the proportion of motorists sticking to the expensive fuel (e) two weeks later was 13 percentage points lower (38% against 51%). This *difference* is likely due to (short-run) consumer “inattention” rather than “tastes:” these 13% of late switchers are likely motorists who might have switched to the cheaper substitute (g) by January 11 but required an additional fortnight to process the relative price change that had already taken place (recall the price path reported in Figure 2). 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” (which we hope to better understand in future research.)

Second, the proportion of motorists choosing the dear fuel over the cheaper one declines with the magnitude of the price difference: compare city-weeks in which the expensive fuel was significantly more expensive—Porto Alegre and Belo Horizonte, Jan 25 with mean p_e/p_g of 91% and 85% respectively, or Curitiba, Mar 29 with p_e/p_g averaging 58%—to city-weeks in which the expensive fuel was moderately more expensive—Curitiba and São Paulo, Jan 25 with mean p_e/p_g of 75% and 74% respectively. This pattern can be explained by either “tastes” or “inattention.”

Finally, notice the lower proportion of motorists choosing the dear fuel (by then g)

in Curitiba, Mar 29 (13%) compared with São Paulo, Mar 29 (26%). By late March, ethanol prices in Curitiba had fallen slightly further than in São Paulo (respectively, a mean p_e/p_g of 57.8% against 59.1%: a “magnitude effect”) and perhaps more gradually (median p_e/p_g over the previous fortnight had fallen 6.7 percentage points in Curitiba against 7.8 percentage points in São Paulo, according to ANP data: an “information diffusion effect”).

Columns I to III of Table 3 report marginal effects of probit specifications estimated off the 1002 observations, where the response variable takes on the value 1 when a motorist chooses the expensive fuel $f = \arg \max_{f \in \{g,e\}} \{p_{g,i}/\widehat{k}_{g,i}, p_{e,i}/\widehat{k}_{e,i}\}$ over the (substantially) cheaper alternative. The dear fuel is g in 289 observations (all collected in March) and e in 713 observations (January). Controls for gender, age, education and car price are insignificant. The result that educational category is not significant may owe to the variance in the quality of the country’s higher education, as well as the hypothesis that not only price misconception but also (unobserved) strong tastes are driving the choice for the dear fuel.

Translating each order into fuel km purchased on the observed fueling occasion (i.e., $q_{f,i}\widehat{k}_{f,i}$), we find that every additional 100 km purchased (the sample mean is 190 km) is associated with an 8% lower probability that the dear fuel is chosen. It is plausible that motorists with greater stakes in the choice of fuel are better informed about relative prices, and/or place less weight on non-price characteristics.

We obtain that motorists who are conscious of habit playing a role in their fueling decisions are significantly more likely to choose the expensive fuel relative to motorists who do not spontaneously invoke habit.

In light of the city-week-level patterns just discussed, the three final variables in specification I measure how expensive the expensive fuel is relative to the cheaper substitute (the “magnitude effect”) and how much more expensive the expensive fuel has become relative to the cheaper substitute over the preceding fortnight (the “information diffusion effect”). First, variable “price distance e above g ” is defined as $\max\{p_{e,i}/k_{e,i} - p_{g,i}/k_{g,i}, 0\} \geq 0$ (with $k_{f,i} = \widehat{k}_{f,i}$); it is strictly positive for the 713 dear-ethanol observations, with the ethanol premium ranging from 0.024 to 0.146 R\$/km). Second, “price distance e below g ” is defined as $\max\{p_{g,i}/k_{g,i} - p_{e,i}/k_{e,i}, 0\} \geq 0$; the gasoline premium in the 289 dear-gasoline observations has a lower range, from 0.024 to 0.061 R\$/km. For perspective, the cheaper fuel across the 1002 observations is priced, on average, at 0.236 R\$/km, so the price premia are considerable. Third, variable “2-week variation in price distance” is defined as $|(p_{e,\cdot} - L2(p_{e,\cdot}))/k_{e,i} - (p_{g,\cdot} - L2(p_{g,\cdot}))/k_{g,i}|$, where $p_{f,\cdot}$ denotes the city-week specific median per-liter price of fuel f (from ANP) and $L2(\cdot)$ denotes the two-week lag operator. For perspective, this price-distance-variation variable 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.

Consistent with the city-week patterns documented above, we find that the likelihood that the dear fuel is chosen over the cheaper one: (i) declines with the magnitude of the price difference (and significantly more so when the cheap fuel is e), and (ii) rises with the recent rate of change of this price difference. (To make sense of marginal effects, divide by 100 for .01 R\$/km changes.)

Relative to column I, column II adds city fixed effects. The effects on the two price-distance variables, still included in the specification, are now identified off only within-city price dispersion (one station versus another) and same-city temporal variation (multiple weeks for São Paulo and Curitiba), but no longer across cities, and as

such become weaker. Similarly, the effect on the recent-price-distance-variation variable is identified off only same-city temporal variation and, though significant, is weaker. Other effects are robust.

Column III replaces city fixed effects (and the city-week-level variation in price distance) by city-week fixed effects. The effects on the two price-distance variables, still included in the specification, are now identified off within-city price dispersion alone and, though negative, are insignificant. Consistent with the data patterns above, the p-value of an equality test: (i) between the São Paulo, Jan 11 and São Paulo, Jan 25 fixed effects is 0.059 (i.e., equality is rejected); and (ii) between the São Paulo, Mar 29 and Curitiba, Mar 29 fixed effects is 0.017 (equality is rejected).

The two final columns of Table 3, to be compared with column III, indicate that these findings are robust to alternative restrictions that prices be “sufficiently unequal” as well as motorists’ assumptions about fuel economy (modifying the price-distance regressors). Specification IV uses observations that satisfy $\max\{p_{g,i}k_{e,i}/(p_{e,i}k_{g,i}), p_{e,i}k_{g,i}/(p_{g,i}k_{e,i})\} \geq 1.1$, i.e., per-km fuel prices were within 10% of each other (note that $N = 1011$, rather than the 1002 observations of columns I to III for which $|p_{g,i}/k_{g,i} - p_{e,i}/k_{e,i}| \geq 0.0242$ R\$/km). Specification V restricts observations to those satisfying $|p_{e,i}/p_{g,i} - 0.7| \geq 0.06$, i.e., the per-liter ethanol-to-gasoline price ratio lay at least 6 percentage points away from the media-reported 70% parity threshold ($N = 1003$).

Appendix: For online publication only

C Appendix: Survey design and data

C.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 (including securing a station manager’s authorization to visit, and traffic that can be chaotic), 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 the National Oil Agency (ANP)’s large retail fueling station database, restricting stations to be branded and to have “purchase invoices” on file (exceptions were Belo Horizonte and Porto Alegre, to make the list long enough). Our intent was to control for the likelihood of product fraud as perceived by consumers across the stations in our study, i.e., the addition of solvents to gasoline and of water to ethanol.⁴² (We flagged any exceptional—e.g., unbranded—stations as such and asked that CNP avoid visiting

⁴²In the Agency’s data, it is noteworthy that: (i) fuel prices are lower at unbranded stations than at branded ones, and (ii) ethanol is even cheaper than gasoline at unbranded stations compared to branded ones. Consider the 5171 observations (station-weeks) in the Agency’s database in the six cities over 7 weeks in Dec-2009 and Jan-2010. Regressing the p_e/p_g ratio on a branded station dummy (71% of this sample is branded) and a full set of city-week effects yields a coefficient on the branded dummy of +1.54% (standard error 0.08%). This is consistent with a higher rate of fraudulent ethanol relative to fraudulent gasoline at unbranded stations.

them, which they did.) The Agency’s database, and thus our list of candidate stations, did not inform on the availability of midgrade and premium gasoline varieties.

The design of a station visit On arriving at a station, our (i.e., 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 gasoline \bar{g} , premium gasoline \check{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 motorists 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 motorist’s car had to meet two qualifying criteria: (i) it was originally manufactured as an FFV (i.e., it was neither a single-fuel car, nor was it subsequently retrofitted);⁴³ and (ii) the FFV was driven for private use (i.e., the car was not a cab or driven at a company’s service).⁴⁴ 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,⁴⁵ 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 (usually via separate nozzles), 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

⁴³Though infrequent, we decided to control for the retrofit possibility by excluding it. Data on the (presumably higher-variance) fuel economy of such modified engines when running on either of the substitute fuels are not available.

⁴⁴We 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.

⁴⁵We base this statement also on dry runs we personally conducted in stations in São Paulo and in Rio de Janeiro. (In practice, 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.)

station (see below).

An observation Each observation was completed in two stages: (i) passively observing the FFV motorist’s choice of fuel (in practice discreetly watching as the motorist placed his order with the serviceman), and (ii) actively interviewing the motorist while his car 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) motorist’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 motorist’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 motorist’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 motorist’s response (e.g., that the consumer was motivated by price), or enter the response verbatim in a sixth option labeled “Other”, which we might later reclassify (see below). A seventh option 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.”

Finally, the representative would enter further vehicle and motorist 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 motorist’s gender: the representative would tick one of two options;

Q8 (Stated) (iii) The motorist’s age and educational attainment: the representative would tick the corresponding category, such as “Over 65 years” and “College, Complete”; and (iv) the motorist’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 motorists would oblige.⁴⁶ 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 duration of a station visit turned out to be 2.5 hours (minimum and maximum durations were respectively 50 minutes and 5hrs25mins): given 12 observations per visit and an expected interview length of under 5 minutes (informed by CNP and supported by our personal dry runs), it appears that our ideal scenario largely prevailed during the course of the survey.

Descriptive statistics of the sample Table 4 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 prices p_g , per-liter ethanol prices p_e , and the price ratio p_e/p_g , at all 180 stations visited (in contrast, midgrade and premium gasoline varieties \bar{g} and \check{g} were available at 164 and 20 visits, respectively, as mentioned in the text). There was considerable cross-sectional and time-series variation in p_e/p_g ; this ratio is summarized for each of the 9 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 hours (stuck in traffic or taking time off before returning to work that same day). On average, less than 3 motorists per station visit refused to take questions (their choices were not recorded).

Table 5 summarizes data that varies across motorists. 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) \bar{g} (and this share rises slightly to 10% if we condition on the midgrade variety being available at the station). (\check{g} was chosen by a mere 0.1% of sampled motorists, which corresponds to a still low 2% of choices in the 20 stations where \check{g} was available.) 2% of motorists purchased a “combo” of g and e , i.e., for which the serviceman needs to handle two nozzles.⁴⁷ Other combos were negligible. The mean fueling ticket was R\$ 47 (R\$ 24

⁴⁶Only 8% of motorists (among 2160) in the sample declined to provide a phone number. (We subsequently based the follow-on phone interviews on the 1991 motorists who left a trail, as detailed in the text.) 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 motorists have missing contact numbers.

⁴⁷Recall that gasoline varieties retailed in Brazil already contained a 20% to 25% proportion of pure ethanol by volume (this was mandated at 25% between Jul-2007 and Jan-2010, and at 20% from Feb-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\$

of g , R\$ 17 of e , R\$ 6 of \bar{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 (not reported in the table) of 22 liters. 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 (also not reported in the table) is once every 7 days.

On the dummy variables indicating “the main reason” invoked spontaneously by a motorist 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 5 exceeds 1 since 17% of motorists were recorded by representatives as volunteering more than one reason (despite the question’s wording, we instructed the representative to record as many reasons as stated by the motorist).

Second, with the wisdom of hindsight, it is now clear 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” could be confused with the top-cited price motivation. (To support this, three-quarters of motorists for whom this reason was ticked purchased less than 28 liters, or a volume lower than 54% of their vehicle’s tank capacity: otherwise, why would a motorist concerned with maximizing range, and thus prone to choosing g thanks to its higher energy content per unit volume, not fill the tank?)

Third, upon inspection we have reclassified 74 of the 172 statements originally marked under “Other” by representatives (who, if unsure as to how to categorize a motorist’s statement, were also instructed to enter this “other” reason verbatim). For example, we have interpreted 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 price characteristics, and reclassified these statements accordingly. Statements remaining in **Other** in the table below are either orthogonal to the precoded menu of five reasons (over half, or 56 to be precise, were motorists invoking “habit”, “accustomed” or a similar reason), or ambiguous or uninformative (e.g., “cost benefit”, “family recommendation” and “my option”):⁴⁸

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

Validating collected data Vehicle model and engine size—for brevity Table 5 shows only vehicle make—required a small amount of cleaning and imputing. The recorded

10, or about 4 liters, combo observations cannot be attributed to the 1-liter startup tank.)

⁴⁸The 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 motorist.

vehicle characteristics reassure us that the qualifying criterion “originally-manufactured FFV” was met, since in the raw data there was only a single observation (among 2160) 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). Also, there was a single observation for which by 2010 Qtr 1 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 car fleet circulating at the end of 2009 from the autoparts industry trade association (Sindipeças). This is available by make, model (including year) and fuel type but, unfortunately, not at the regional level. We then compared the proportion of car models in our six-city survey sample to those in the nationwide stock of FFVs (only 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 (13.3% of our observations vs. 11.5% of the stock, in all its flex versions), undersample the VW Gol (10.1% vs. 13.2%), and sample the GM Corsa at the nationally representative rate (7.7% vs 7.8%). With hypothetical data on cars circulating in each of the sampled cities (or neighborhoods) by time of day, one could formally test the distribution of car models in our survey records. In the absence of such data, we did not detect any anomaly.⁴⁹

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—see the bottom rows of Table 5, recalling that the median motorist is male and middle-aged—an old survey by São Paulo city’s traffic authority (CET 1987) found that 84% of the city’s weekday motorists were men; 13% of motorists were under 25 years, 52% were 25-40y, 31% were 40-60y, and 4% were 60y+. That half of our sample declared having completed college education could owe in part to cultural factors, but one should note that subjects belong to Brazil’s better-off households, as they had a relatively new car at their personal disposal (the average age of FFVs in circulation was 2.6 years according to the Sindipeças data). Also, quality varies significantly in the country’s provision of higher education.

We also validated several station-level fuel prices in our sample, as recorded by our representatives, against prices (g and e only) in the National Oil Agency’s database.

C.2 Other data sources

Vehicle-specific fuel economy As mentioned in the text, we use laboratory measurements endorsed by Inmetro to predict the fuel economy of the FFVs in our sample. Following U.S. EPA guidelines, Inmetro reports the kilometerage per liter (kpl) for re-tailed fuels g (E22) and e (E100) under separate driving cycles—urban and highway—for a fairly broad selection of new FFVs. Note that “cleaning” additives and higher octane rating in upmarket gasoline varieties \bar{g} and \check{g} do not materially affect their kpl relative to the regular variety g , and that e (known as “hydrated” ethanol) already includes a residual water content, as re-tailed. Specifically, lab measures, first published in 2009, were available on 23 2009 models and on 44 2010 models. Since measures were not

⁴⁹The distribution of car models varies significantly across cities in our survey sample. Unsurprisingly, a χ^2 goodness of fit test indicates that car model composition in our (unweighted, six-city) sample differs significantly from that in the nationwide stock, though differences do not appear to be large.

available for several FFVs in our survey sample, we took an indirect route. Using the available lab data, we first projected kpl on other vehicle characteristics; we then used the fitted regression to predict kpl for each vehicle in our survey sample based on its observed characteristics. One might worry that surveyed cars, unlike those tested in the lab, were not necessarily new. However, important for our purpose of analyzing fuel choice, we can plausibly assume that the *ratio* of kpl for the two fuels, $k_{e,i}/k_{g,i}$, does not vary systematically with car use, condition or age (recall that surveyed cars were manufactured no earlier than 2003).

Table 6 presents these auxiliary regressions. We predict six vehicle-specific fuel economy variables, respectively labeled in columns I to VI k_e/k_g (i.e., urban driving), $k_e^{highway}/k_g^{highway}$, k_g , k_e , $k_g^{highway}$ and $k_e^{highway}$ (we add a superscript to denote highway as opposed to urban driving). 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 predict fuel economy using: (a) specification “a,” which includes car model fixed effects, for those car models in our survey sample which were tested in the lab (i.e., for which there is a model fixed effect, e.g., the GM Celta 1.0 liter)—this is the case for 66% of FFVs in our survey sample; and (b) specification “b,” which includes carmaker fixed effects and car segment fixed effects, for the remaining car models—34% of our survey sample—which were not tested in the lab. In this latter case, we need to use lab data on similar car 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 the kpl ratio, k_e^{city}/k_g^{city} , evaluated at the mean of the regressors (in the lab sample) is 67.7% (and 68.1% excluding Renault models, see below), with robust standard error of 0.1% or 0.2%. The kpl ratio is similar across FFVs: the interdecile range (p90-p10) is 3.3% (i.e., 69.4% – 66.1%). This indicates that the 70% relative price parity threshold that is repeatedly reported in the media is not wide off the mark.⁵⁰ (In the data, Renault is an outlier: the firm produced all 5 models with kpl ratio below 65%, i.e., low ethanol economy relative to gasoline in comparison to other makes. Having entered Brazil only a decade earlier, Renault had less than a 5% market share in 2010.)

Vehicle-specific prices and tank capacities We matched the vehicles in our survey to an external dataset on used car prices.⁵¹ Based on observed make and model (we do not observe the surveyed car’s age), we took the price in the secondary market in February 2010 for the most recent year-version (“model year”) of that make and model. For example, Fiat was still producing its Palio 1.0 (liter) in early 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-tenths of our observations with 2009 or 2010 model prices. The mean price across the 2160 FFVs in our sample is R\$ 33,602 (minimum price is R\$

⁵⁰The mean predicted value for surveyed vehicles (adjusting for slightly varying blended gasoline composition between Jan-2010 and Mar-2010—recall note 15) is 68.7% (standard deviation 1.6%), not shown in Table 6.

⁵¹We accessed the QuatroRodas/Molicar/Fipe price tables on March 9-11, 2010; these are available at <http://quatrorodas.abril.com.br/compre-seu-carro/tabela-de-precos>.

21,216, maximum price is R\$ 90,732).

It is reassuring that surveyed cars fueling at stations that carried midgrade gasoline on top of regular (presumably located in wealthier neighborhoods) display a higher mean price than cars observed in stations carrying only the regular gasoline variety; the p-value of a test of the equality of means is 0.031 (a mean car price of R\$ 33,780 in stations where \bar{g} was available against R\$ 31,784 in the fewer stations where it was not). Similarly, surveyed cars fueling at stations that carried the even more upmarket \check{g} (besides \bar{g} and g) have a higher mean price than cars at stations carrying only \bar{g} and g (the equality test has p-value of 0.034).

As might be expected, cars sampled in relatively less affluent Recife, in the country’s Northeast region, have a lower mean price (R\$ 30,834) than cars sampled in the five southern/southeastern cities (a mean price of R\$ 33,948; an equality test yields a p-value of 0.000).

Since zip-code-level incomes are not available, we take the combined price of the 12 cars 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 to 0.44 million in less affluent Recife, R\$ 0.34 to 0.57 million in more affluent Rio de Janeiro): some neighborhoods, or motorists along certain routes, are more affluent than others (recall from Figure 3 that we cast the net widely). To support the notion that variation in station-level car 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 visited more than once (in different weeks). The aggregate price of the 12 cars surveyed during each visit for a station visited multiple times varies little compared to cross-station variation. For example, the aggregate car price across the 60 São Paulo station visits varies from R\$ 0.32 to 0.53 million, with standard deviation of R\$ 0.04 million; in contrast, the aggregate car price for a same São Paulo station visited in January and March varies from R\$ 0.37 to 0.40 million (to provide a typical example).

Similarly, we obtained vehicle-specific tank capacities from Carrosnaweb. The mean tank capacity in our survey sample is 51 liters (minimum capacity is 42 liters, maximum capacity is 80 liters).

Appendix: For online publication only

D Appendix: Phone-based interview evidence

The table below details the responses of the $N' = 607$ 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?*”). (As mentioned in the text, each interview was randomly assigned an order in which the two fuels would appear in statements to 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?”). (Motorists 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—for brevity, these responses are omitted.)

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 the text but not in the table).

Favorable views comparative	Motorists observed in station	Total phone respond.	“Environ.”		“Vehicle performance”			“Country”		
			Q1’ → ethanol*	Q2’ → ethanol*	Q2’ → gasoline**	Q3’ → ethanol*				
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%
Subtot. 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%
Subtot. e importers	720	191	149	78%	51	27%	103	54%	116	61%
Total	2160	607	500	82%	187	31%	300	49%	422	70%

Notes: Number and proportion of total phone respondents who selected: *ethanol or **gasoline (other responses are omitted).

The following table evaluates the answers of 403 phone respondents who, as mentioned in the text, 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 $0.7p_{gi}$	28	7%	41%
Reasonably alludes to “cost per km” even if briefly	54	13%	54%
Refers to “use of table” (internet/online calculator/cell phone app, etc)	25	6%	60%
Appears confused with respect to, or does not remember, cost conversion*	39	10%	70%
Motorist states or suggests that he does not do, or know, cost conversion*	86	21%	91%
Avoids directly answering question	35	9%	100%

Notes: *Fuel choice might still be guided by, e.g., the radio, colleagues or the serviceman.

Specification:	I: m.e. (s.e.)	II: m.e. (s.e.)	III: m.e. (s.e.)
Pr (Motorist i chooses fuel e)	Pr = 0.43	Pr = 0.43	Pr = 0.44
Price of e , R\$/km (mean = .268)	-3.80*** (0.62)	-4.43*** (0.69)	-6.39*** (0.51)
Price of g , R\$/km (mean = .246)	3.27*** (0.63)	4.08*** (0.68)	4.73*** (0.55)
Price of \bar{g} , R\$/km (mean = .257)	0.53* (0.31)	0.35 (0.27)	1.66*** (0.33)
1(female) (mean = .342)	-0.01 (0.03)	-0.01 (0.03)	-0.01 (0.03)
1(aged 25 to 40 years) (mean = .463)	-0.07 (0.04)	-0.06 (0.04)	-0.07* (0.04)
1(aged 40 to 65 years) (mean = .395)	-0.07* (0.04)	-0.06 (0.04)	-0.08** (0.04)
1(aged more than 65 years) (mean = .037)	-0.27*** (0.05)	-0.25*** (0.06)	-0.28*** (0.08)
1(some secondary schooling) (mean = .310)	0.04 (0.05)	0.06 (0.05)	0.04 (0.05)
1(some college education) (mean = .618)	0.02 (0.05)	0.04 (0.05)	0.03 (0.05)
1(car usage in upper quartile) (mean = .229)	-0.08*** (0.03)	-0.09*** (0.03)	-0.09*** (0.03)
1(car price in upper quartile) (mean = .262)	-0.05** (0.03)	-0.06** (0.03)	-0.05* (0.03)
Price of 12 cars in station (mean = .403)	0.28 (0.33)	0.17 (0.35)	0.11 (0.31)
1(São Paulo) (Producer, mean = .333)	0.10 (0.15)	0.11 (0.16)	
1(Curitiba) (Producer, mean = .222)	0.16 (0.14)	0.21 (0.15)	
1(Recife) (Producer, mean = .111)	0.08 (0.14)	0.10 (0.15)	
1(Rio de Janeiro) (Importer, mean = .111)	-0.04 (0.14)	-0.00 (0.15)	
1(Belo Horizonte) (Importer, mean = .111)	-0.14 (0.13)	-0.09 (0.15)	
1(Porto Alegre) (Importer, mean = .111)	-0.23* (0.13)	-0.19 (0.15)	
Envir. & e -premium $\geq 10\%$ (mean = .021)		0.51*** (0.03)	
Envir. & g -premium $\geq 10\%$ (mean = .008)		0.27* (0.15)	
Envir. & premium $< 10\%$ (mean = .027)		0.37*** (0.06)	
Engine & e -premium $\geq 10\%$ (mean = .068)		-0.20*** (0.06)	
Engine & g -premium $\geq 10\%$ (mean = .013)		-0.37*** (0.04)	
Engine & premium $< 10\%$ (mean = .042)		-0.24*** (0.05)	
Pr (Motorist i chooses fuel g)	Pr = 0.50	Pr = 0.51	Pr = 0.48
Price of e , R\$/km (mean = .268)	3.27*** (0.63)	4.08*** (0.68)	4.73*** (0.55)
Price of g , R\$/km (mean = .246)	-3.87*** (0.84)	-4.74*** (0.95)	-5.47*** (1.10)
Price of \bar{g} , R\$/km (mean = .257)	0.59 (0.61)	0.66 (0.70)	0.74 (1.23)
(...other marginal effects omitted)			
Pr (Motorist i chooses fuel \bar{g})	Pr = 0.07	Pr = 0.06	Pr = 0.08
Price of e , R\$/km (mean = .268)	0.53* (0.31)	0.35 (0.27)	1.66*** (0.33)
Price of g , R\$/km (mean = .246)	0.59 (0.61)	0.66 (0.70)	0.74 (1.23)
Price of \bar{g} , R\$/km (mean = .257)	-1.12 (0.74)	-1.01 (0.80)	-2.40* (1.29)
(...other marginal effects omitted)			
Number of motorists	2160	2160	2160
Total number of alternatives	6288	6288	6288
Log likelihood	-1697	-1618	-1764
$\sigma_{\bar{g}}$	1.44 (1.08)	2.06 (1.49)	1.37 (1.47)
$\rho_{g,\bar{g}}$ [in III: $\rho_{e,\bar{g}}$]	0.01 (0.83)	0.32 (0.68)	0.56 (0.62)

Table 1: Multinomial probit estimated marginal effects (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 (if available at the station). 1(.) denotes the indicator function. Combined value of 12 cars sampled during each station visit is in million reais. To obtain standard errors on price effects in column III, the free correlation parameter was changed to (ethanol, midgrade gasoline) (while estimated coefficients are robust). Station visit-clustered standard errors in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

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.)
Pr (Motorist i chooses fuel e)	Pr = 0.43	Pr = 0.60	Pr = 0.41	Pr = 0.47	Pr = 0.41	Pr = 0.45
Price of e (IV: R\$/liter, 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)
Price of g (IV: R\$/liter, 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)
Price of \bar{g} (IV: R\$/liter, 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)
1(female)	-0.01 (0.03)	-0.01 (0.06)	0.02 (0.05)	-0.01 (0.06)	0.00 (0.04)	-0.02 (0.04)
1(aged 25 to 40 years)	-0.07* (0.04)	-0.12** (0.06)	-0.08 (0.08)	-0.03 (0.06)	-0.13** (0.06)	-0.02 (0.05)
1(aged 40 to 65 years)	-0.07* (0.04)	-0.10* (0.06)	-0.07 (0.06)	-0.04 (0.08)	-0.14** (0.06)	-0.00 (0.05)
1(aged more than 65 years)	-0.26*** (0.05)	-0.29* (0.15)	-0.28** (0.11)	-0.25** (0.10)	-0.29*** (0.07)	-0.22*** (0.09)
1(some secondary schooling)	0.04 (0.05)	0.07 (0.07)	0.09 (0.08)	0.05 (0.08)	0.01 (0.08)	0.08 (0.07)
1(some college education)	0.02 (0.05)	0.05 (0.05)	0.03 (0.08)	0.07 (0.08)	0.02 (0.07)	0.03 (0.07)
1(car usage in upper quartile)	-0.08*** (0.03)	-0.08* (0.05)			-0.08* (0.04)	-0.08* (0.04)
1(car price in upper quartile)	-0.06** (0.03)	-0.11*** (0.04)				
Price of 12 cars in station	0.20 (0.32)	0.59 (0.62)	-0.05 (0.04)	-0.10 (0.05)		
1(São Paulo)	0.10 (0.16)	0.07 (0.44)	-0.03 (0.43)	0.54 (0.50)	-0.12 (0.40)	0.54 (0.46)
1(Curitiba)	0.16 (0.15)	0.07 (0.28)	0.13 (0.48)	-0.07 (0.23)	0.29 (0.23)	-0.07 (0.22)
1(Recife)	0.10 (0.14)		0.17 (0.40)	0.06 (0.22)	0.31 (0.21)	0.06 (0.22)
1(Rio de Janeiro)	-0.01 (0.15)		0.14 (0.26)	-0.10 (0.21)	0.23 (0.20)	-0.03 (0.20)
1(Belo Horizonte)	-0.10 (0.14)		0.02 (0.28)	-0.21 (0.22)	0.13 (0.23)	-0.16 (0.19)
1(Porto Alegre)	-0.22 (0.15)		-0.03 (0.29)	-0.39 (0.19)	0.07 (0.22)	-0.29 (0.16)
Past 2-wk increase in ethanol price		2.28 (2.39)	-0.23 (0.23)	-0.36 (0.16)	-0.09 (0.25)	-0.34 (0.18)
Pr (Motorist i chooses fuel g)	Pr = 0.50	Pr = 0.36	Pr = 0.52	Pr = 0.46	Pr = 0.50	Pr = 0.50
Price of e (IV: R\$/liter, 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)
Price of g (IV: R\$/liter, 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)
Price of \bar{g} (IV: R\$/liter, 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)
Number of motorists	2160	1200	1063	772	1100	1060
Total number of alternatives	6288	3528	3112	2240	3214	3074
Log likelihood	-1691	-917	-835	-604	-888	-796
$\sigma_{\bar{g}}$	1.78 (1.22)	1.81 (0.93)	2.67 (1.81)	0.43 (1.61)	1.58 (1.31)	1.51 (1.59)
$\rho_{g,\bar{g}}$	0.27 (0.94)	0.66 (2.04)	0.97 (0.24)	0.74 (2.30)	0.64 (0.76)	-0.24 (1.02)

Table 2: Multinomial probit estimated marginal effects (at the sample mean) for alternative specifications. Some effects are not reported (see text and the notes to Table 1). Station visit-clustered standard errors in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$

Specification	I	II	III	IV	V
1(female)	-0.021	-0.015	-0.016	-0.007	0.016
mean (I to III) = .357	(0.031)	(0.031)	(0.031)	(0.032)	(0.032)
1(aged 25 to 40 years)	0.083	0.069	0.075	0.057	0.090*
mean = .457	(0.053)	(0.053)	(0.054)	(0.054)	(0.053)
1(aged 40 to 65 years)	0.048	0.030	0.035	0.018	0.075
mean = .399	(0.053)	(0.055)	(0.056)	(0.054)	(0.052)
1(aged more than 65 years)	0.044	0.023	0.029	0.012	0.090
mean = .038	(0.098)	(0.095)	(0.098)	(0.099)	(0.117)
1(some secondary schooling)	0.064	0.050	0.051	0.063	0.031
mean = .326	(0.062)	(0.063)	(0.062)	(0.063)	(0.061)
1(some college education)	0.036	0.011	0.021	0.010	-0.011
mean = .593	(0.056)	(0.058)	(0.057)	(0.056)	(0.057)
1(car price in upper quartile)	0.028	0.020	0.022	0.018	0.048
mean = .276	(0.036)	(0.037)	(0.037)	(0.035)	(0.034)
Price of 12 cars in station	-0.121	-0.032	-0.071	0.058	-0.142
mean = .408	(0.372)	(0.386)	(0.381)	(0.364)	(0.350)
Km purchased on occasion (x100)	-0.078***	-0.077***	-0.074***	-0.056***	-0.051***
mean = 1.898	(0.018)	(0.018)	(0.018)	(0.018)	(0.017)
1(justifies choice through habit)	0.424***	0.413***	0.432***	0.378***	0.305***
mean = .026	(0.081)	(0.084)	(0.082)	(0.089)	(0.099)
Price distance e above g	-4.106***	0.255	-0.332	-0.001	-0.701
mean = .035	(1.158)	(1.526)	(1.604)	(0.005)	(0.787)
Price distance e below g	-12.409***	-7.526***	-2.348	-0.004	-0.695
mean = .010	(1.653)	(1.834)	(4.044)	(0.005)	(0.878)
2-week variation in price distance	10.176***	5.510**			
mean = .014	(2.510)	(2.777)			
1(São Paulo, multiple weeks)		-0.038			
1(Curitiba, multiple weeks)		-0.073			
1(Recife, Jan 25)		0.115	0.237	0.156	0.200
1(Rio de Janeiro, Jan 25)		-0.119	-0.040	-0.101	-0.010
1(Belo Horizonte, Jan 25)		-0.203*	-0.151	-0.198*	-0.088
1(Porto Alegre, Jan 25)		-0.260***	-0.241***	-0.265***	-0.184
1(São Paulo, Jan 11)			0.113	0.020	0.186
1(São Paulo, Jan 25)			-0.039	-0.109	-0.079
1(São Paulo, Mar 29)			-0.035	-0.121	-0.093
1(Curitiba, Jan 25)			0.078	0.005	0.184
1(Curitiba, Mar 29)			-0.166	-0.190*	-0.134
Number of observations	1002	1002	1002	1011	1003
Log likelihood	-524.9	-511.4	-507.8	-517.0	-501.9

Table 3: Probit estimated marginal effects (at the sample mean). Success is "Expensive fuel is chosen over the cheaper substitute." An observation is a motorist in the restricted sample facing sufficiently unequal prices across regular gasoline and ethanol. 1(.) denotes the indicator function. See the text for exact variable definitions. Station visit-clustered standard errors in parentheses. For city and city-week fixed effects, only significance levels are shown due to space. * $p < .1$, ** $p < .05$, *** $p < .01$

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
p_g (R\$/liter)...	180	2.515	.146	2.199	2.989
...for city-weeks in January	140	2.550	.125	2.199	2.989
...for city-weeks in March	40	2.391	.148	2.229	2.989
p_e (R\$/liter)...	180	1.879	.316	1.239	2.549
...for city-weeks in January	140	2.017	.196	1.690	2.549
...for city-weeks in March	40	1.399	.129	1.239	1.899
p_e/p_g (ratio)...	180	.745	.105	.520	1.050
...for São Paulo, week of Jan 11	20	.742	.027	.692	.800
...for São Paulo, week of Jan 25	20	.747	.032	.692	.833
...for Curitiba, week of Jan 25	20	.747	.021	.692	.775
...for Recife, week of Jan 25	20	.745	.022	.698	.784
...for Rio de Janeiro, week of Jan 25	20	.805	.037	.739	.873
...for Belo Horizonte, week of Jan 25	20	.846	.026	.800	.885
...for Porto Alegre, week of Jan 25	20	.905	.041	.858	1.050
...for São Paulo, week of Mar 29	20	.591	.038	.520	.667
...for Curitiba, week of Mar 29	20	.578	.021	.541	.609
$p_{\bar{g}}$ (R\$/liter)	164	2.618	.185	2.25	3.499
$p_{\bar{g}}/p_g$ (ratio)	164	1.039	.030	1	1.171
$p_{\check{g}}$ (R\$/liter)	20	3.213	.240	2.799	3.699
$p_{\check{g}}/p_{\bar{g}}$ (ratio)	20	1.162	.093	1.034	1.375
$p_{\check{g}}/p_g$ (ratio)	20	1.223	.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
...for São Paulo or Curitiba, week of Jan 25	40	4.875	2.053	2	12
...for São Paulo or Curitiba, week of Mar 29	40	5.175	2.591	2	16
Number of e -nozzles...	180	3.900	2.302	1	16
...for São Paulo or Curitiba, week of Jan 25	40	3.950	2.183	1	10
...for São Paulo or Curitiba, week of Mar 29	40	4.500	2.792	1	16
Number of \bar{g} -nozzles if \bar{g} is available	164	3.677	2.419	1	16
Number of \check{g} -nozzles if \check{g} is available	20	2.050	.510	1	4
1(fueling station is branded)	180	.989	.105	0	1
1(fueling station is BR branded)	180	.294	.457	0	1
1(fueling station is Shell branded)	180	.267	.443	0	1
1(fueling station is Ipiranga branded)	180	.189	.393	0	1
1(station visit started peak hour, weekday)	180	.317	.466	0	1
1(station visit started non-peak hour, weekday)	180	.467	.500	0	1
1(station visit on Saturday)	180	.217	.413	0	1
Duration of station visit (hours)	180	2.543	.971	.833	5.417
Representative's travel time between stations	100	1.804	1.448	.133	7.75
Number of motorists who refused to participate	180	2.844	4.026	0	22

Table 4: Summary statistics: Data directly collected in, or constructed from, station-level forms. An observation is a station visit. 1(.) denotes the indicator function.

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Purchased (only) g ...	2160	.451	.498	0	1
...for city-weeks in January	1680	.520	.500	0	1
Purchased (only) e ...	2160	.435	.496	0	1
...for city-weeks in March	480	.719	.450	0	1
Purchased (only) \bar{g}	2160	.087	.282	0	1
Purchased (only) \check{g}	2160	.001	.037	0	1
Purchased a “combo” of g and e	2160	.024	.152	0	1
Purchased any other “combo”	2160	.001	.037	0	1
Value of fuel purchased (in R\$)	2160	46.973	29.601	10	158
Value of g purchased (in R\$)...	2160	24.052	33.354	0	150
...if purchasing (only) g	975	51.932	31.659	10	150
Value of e purchased (in R\$)...	2160	17.081	23.877	0	140
...if purchasing (only) e	940	37.459	22.401	10	140
Value of \bar{g} purchased (in R\$)...	2160	5.752	21.126	0	150
...if purchasing (only) \bar{g}	188	65.477	34.472	10	150
1(main reason is price)	2160	.683	.465	0	1
1(main reason is range or price)	2160	.263	.440	0	1
1(main reason is environment)	2160	.056	.230	0	1
1(main reason is engine performance)	2160	.047	.211	0	1
1(main reason is engine startup)	2160	.083	.276	0	1
1(other main reason)	2160	.045	.208	0	1
1(does not know the main reason)	2160	.003	.057	0	1
On last two stops, twice chose g ...	2160	.479	.500	0	1
...if (now) purchasing (only) g	975	.825	.380	0	1
On last two stops, twice chose e ...	2160	.372	.484	0	1
...if (now) purchasing (only) e	940	.778	.416	0	1
On last two stops, alternated e, g	2160	.149	.356	0	1
Of last two stops, twice at this station	2160	.513	.500	0	1
Of last two stops, once at this station	2160	.219	.413	0	1
Of last two stops, none at this station	2160	.269	.444	0	1
Car usage (km/week)	1835	296.094	319.930	5	3500
1(vehicle make is Fiat)	2160	.281	.449	0	1
1(vehicle make is General Motors)	2160	.248	.432	0	1
1(vehicle make is Volkswagen)	2160	.231	.422	0	1
1(vehicle make is Ford)	2160	.102	.303	0	1
1(male)	2160	.658	.475	0	1
1(aged 25 to 40 years)	2160	.463	.499	0	1
1(aged 40 to 65 years)	2160	.395	.489	0	1
1(aged more than 65 years)	2160	.037	.188	0	1
1(completed primary school)	2160	.059	.235	0	1
1(dropped out of secondary school)	2160	.026	.160	0	1
1(completed secondary school)	2160	.284	.451	0	1
1(reached college, incomplete)	2160	.121	.326	0	1
1(completed college)	2160	.497	.500	0	1
1(provided a contact number)	2160	.922	.269	0	1

Table 5: Summary statistics: Data directly collected in, or constructed from, motorist-level forms. An observation is an FFV-qualifying motorist. 1(.) denotes the indicator function.

Dependent variable:	$\frac{k_e}{k_g}$	$\frac{k_e^{highway}}{k_g^{highway}}$	k_g	k_e	$k_g^{highway}$	$k_e^{highway}$
Specification:	Ia	IIa	IIIa	IVa	Va	VIa
Size of the engine (liters)	-0.011** (0.005)	-0.004 (0.009)	-1.036*** (0.233)	-0.814*** (0.158)	-0.576* (0.312)	-0.430** (0.184)
1(non-manual transmission)	-0.013 (0.008)	-0.007 (0.009)	-0.015 (0.167)	-0.147 (0.104)	0.251 (0.216)	0.072 (0.142)
Intercept	0.709*** (0.009)	0.684*** (0.013)	11.250*** (0.388)	7.940*** (0.282)	12.156*** (0.438)	8.302*** (0.257)
Car 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 (0.002)	0.674 (0.002)	10.393 (0.043)	7.036 (0.028)	12.652 (0.057)	8.528 (0.033)
Specification:	Ib	IIb	IIIb	IVb	Vb	VIb
Size of the engine (liters)	-0.012** (0.005)	-0.003 (0.008)	-1.530*** (0.330)	-1.170*** (0.239)	-0.956** (0.463)	-0.685** (0.337)
1(non-manual transmission)	-0.013** (0.006)	-0.006 (0.007)	0.013 (0.147)	-0.122 (0.098)	0.145 (0.222)	0.016 (0.148)
1(car segment is compact)	-0.007* (0.004)	-0.006 (0.006)	-0.553* (0.295)	-0.454** (0.208)	-0.536 (0.380)	-0.431 (0.262)
1(car segment is midsize)	-0.007 (0.004)	-0.004 (0.007)	-0.263 (0.295)	-0.257 (0.216)	-0.447 (0.421)	-0.350 (0.293)
1(car segment is full-size)	0.002 (0.015)	-0.003 (0.015)	-0.925** (0.371)	-0.594** (0.253)	0.138 (0.534)	0.064 (0.372)
1(car segment is small truck)	-0.011** (0.005)	-0.004 (0.008)	-0.659 (0.455)	-0.571* (0.307)	-1.656** (0.737)	-1.174** (0.542)
Intercept	0.710*** (0.008)	0.689*** (0.011)	12.685*** (0.580)	8.945*** (0.416)	13.636*** (0.759)	9.369*** (0.529)
Carmaker 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 (0.001)	0.674 (0.002)	10.393 (0.070)	7.036 (0.048)	12.652 (0.110)	8.528 (0.078)
Number of observations	67	67	67	67	67	67

Table 6: Auxiliary OLS regressions for predicting vehicle-specific fuel economy based on laboratory test data reported by Inmetro. An observation is a car model in the Inmetro test sample. In "a" specifications, car model fixed effects are included (the Fiat Palio subcompact dummy variable is omitted). In "b" specifications, carmaker fixed effects and car segment fixed effects are included (the Fiat and the subcompact dummy variables are omitted). 1(.) denotes the indicator function. Heteroscedasticity-robust standard errors in parentheses. * p<.1, ** p<.05, *** p<.01

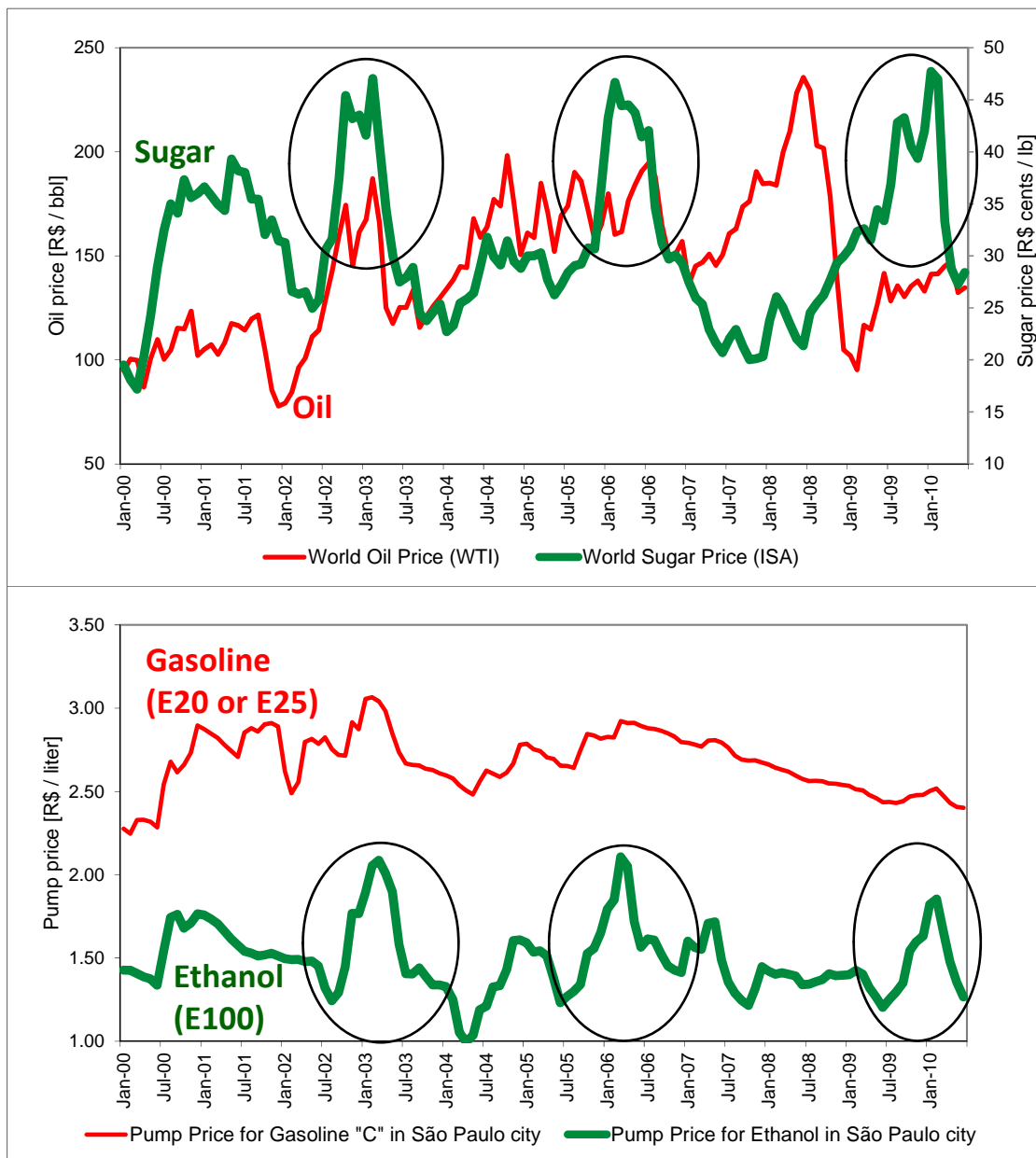


Figure 1: The pump price of ethanol peaks when the world price of sugar peaks. Upper panel: World sugar price (ISA, R\$ cents per pound) and World oil price (WTI, R\$ per barrel). Lower panel: Ethanol (E100) and Gasoline (regular, E20-25) prices at the pump in the city of São Paulo (R\$ per liter). All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

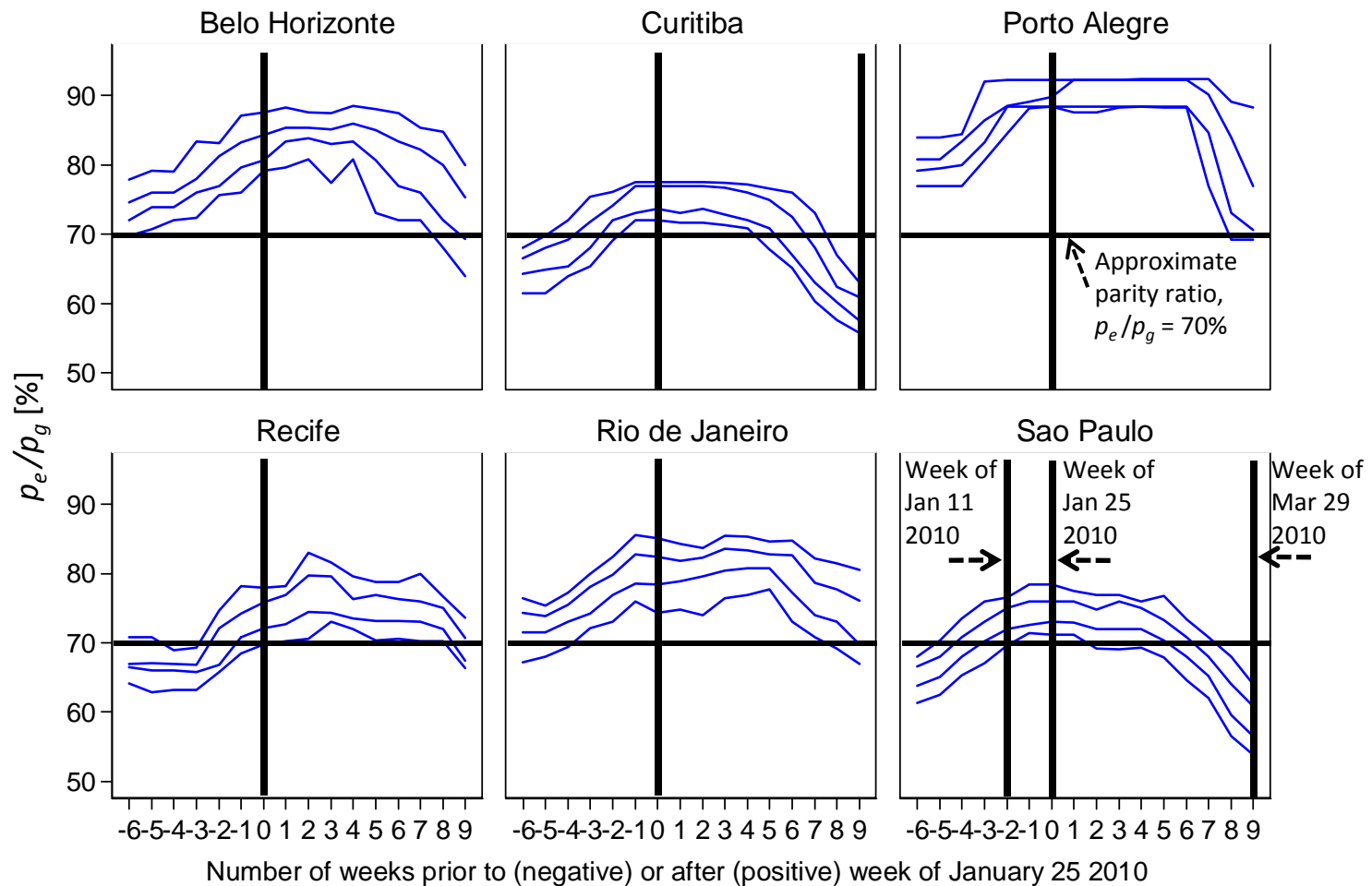


Figure 2: Opportune price variation in the 1st Quarter of 2010. Percentiles (5th, 25th, 75th, and 95th) of the distribution across branded stations of the ethanol-to-regular-gasoline per-liter price ratio, in six cities in each of several weeks running up to and following the week of January 25, 2010. The 5th, 25th, 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.

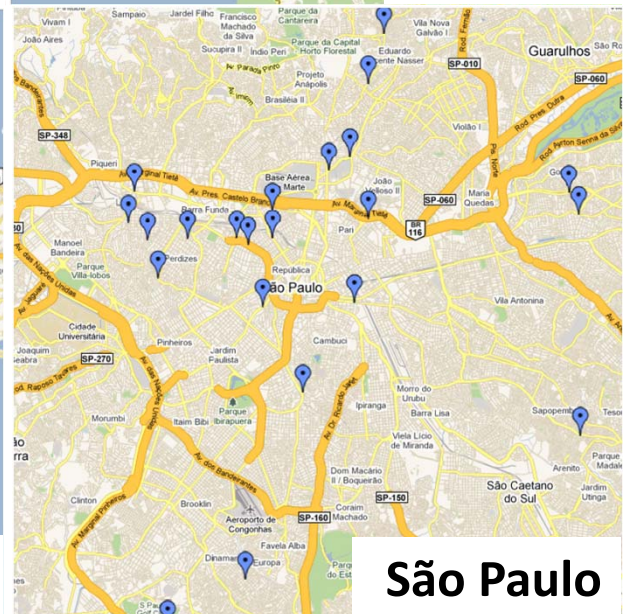
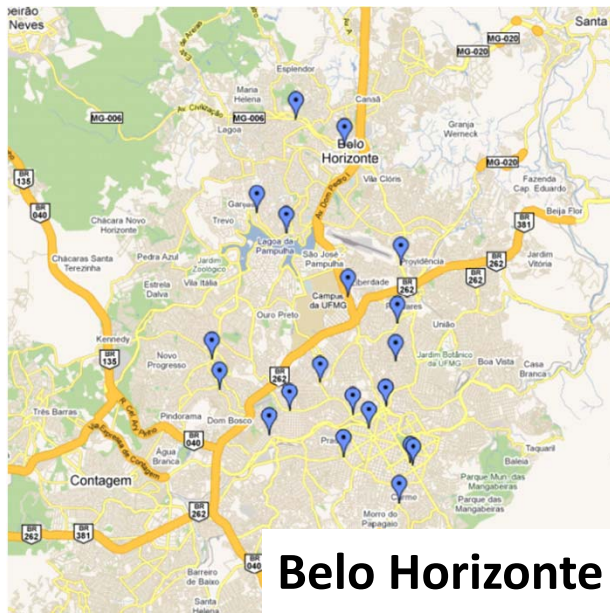


Figure 3: Location of retail fueling stations in our sample. Map source: Google.

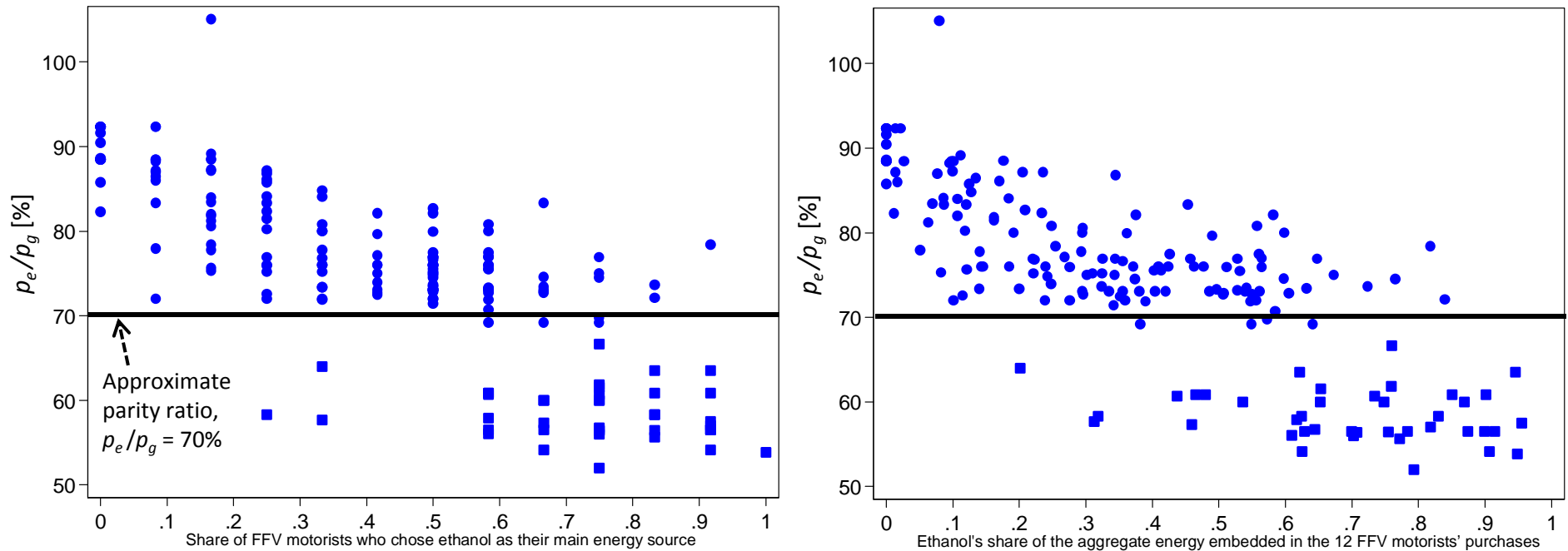


Figure 4: Empirical demand by FFV motorists at the station level. Per-liter ethanol price relative to regular gasoline (p_e/p_g) plotted against ethanol's overall "share" in the 12 choices observed in each station visit. The left and right panels respectively plot the "unweighted" and "weighted" ethanol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

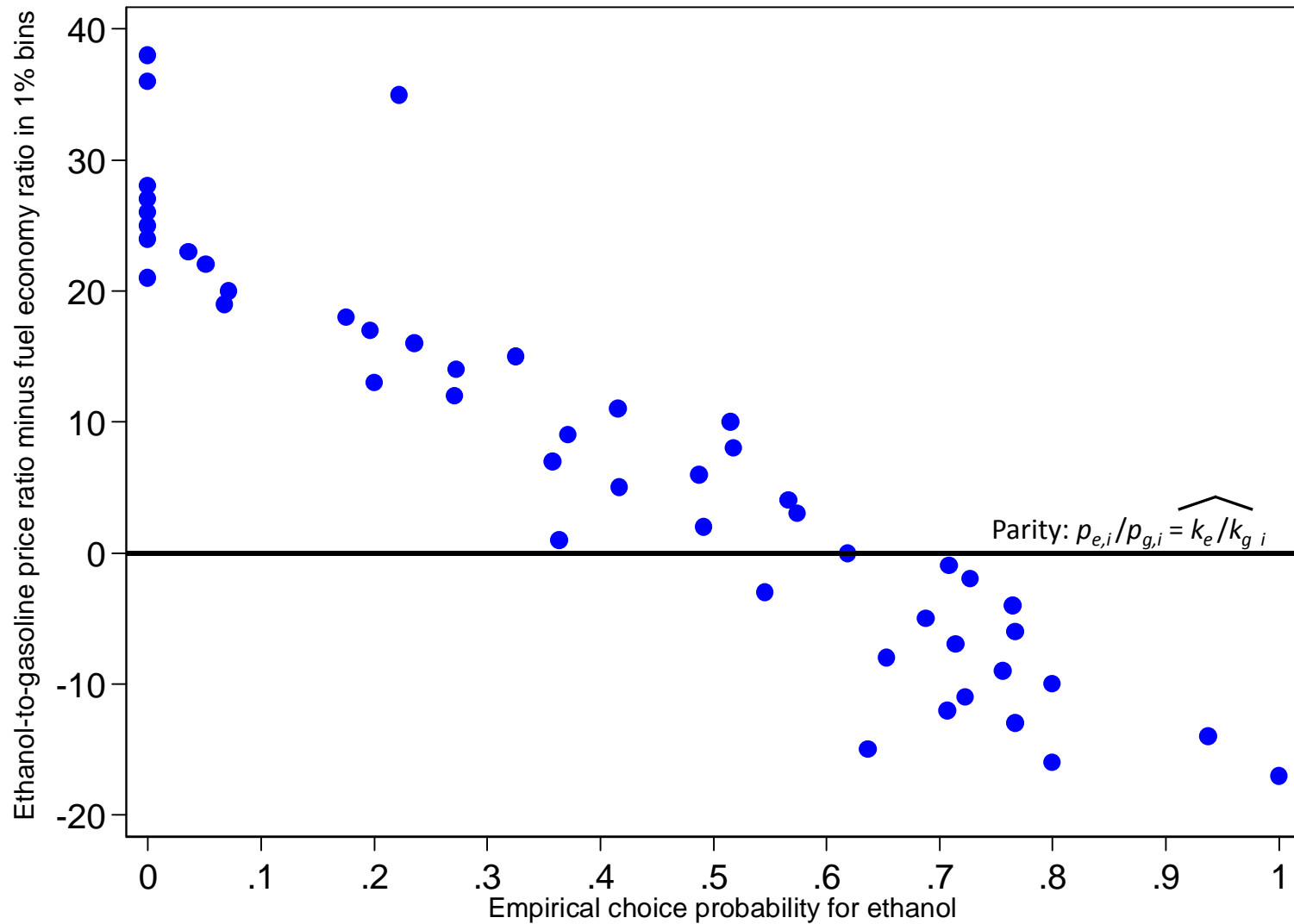


Figure 5: Empirical choice probability for ethanol controlling for energy-adjusted relative prices. An observation is an integer category defined over 2160 choices. The vertical axis shows 1 ppt “bins” for the difference between the ethanol-to-regular-gasoline per-liter price ratio the motorist faced at the station and the predicted ethanol-to-gasoline kpl ratio for his FFV. The horizontal axis reports the proportion of motorists in that bin who chose ethanol as their dominant energy source. (See text.) Source: Own survey.

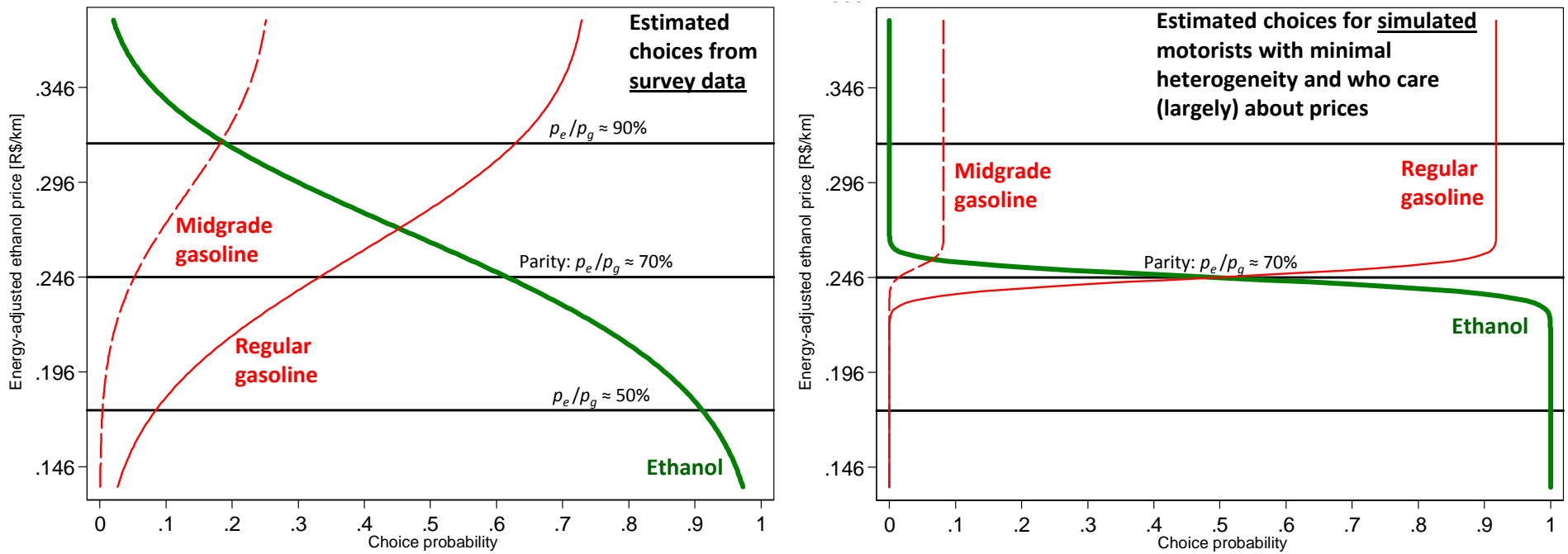


Figure 6: Illustrating unobserved consumer heterogeneity. Left panel: A “median” motorist’s fuel choice probabilities estimated from the survey data. Source: Specification III estimates, to conservatively reduce the price range over which substitution takes place. Right panel: Fuel choices probabilities estimated for (2160) simulated motorists 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 motorist’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 kpl ratio of 70%).

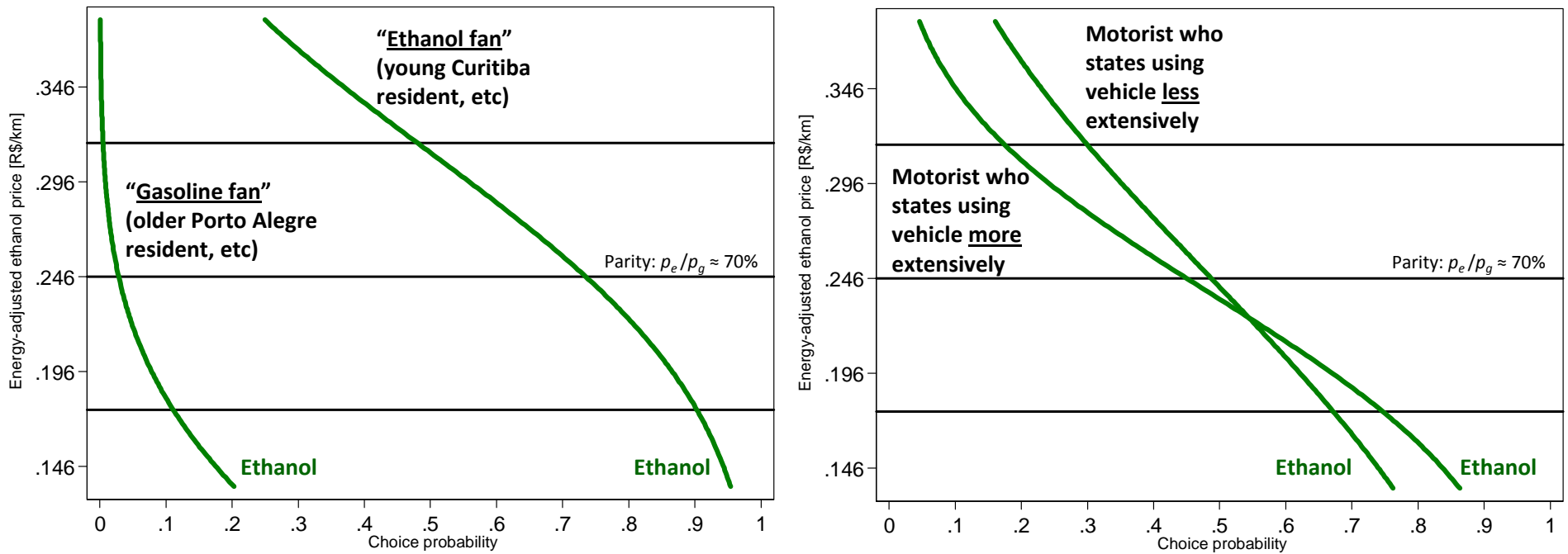


Figure 7: Illustrating observed consumer heterogeneity. Left panel: Ethanol choice probabilities for two hypothetical polar consumers, an “ethanol fan” and a “gasoline fan.” Source: Specification I estimates. Right panel: Ethanol choices probabilities for consumers with above-median versus below-median 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 motorist’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 kpl ratio of 70%).

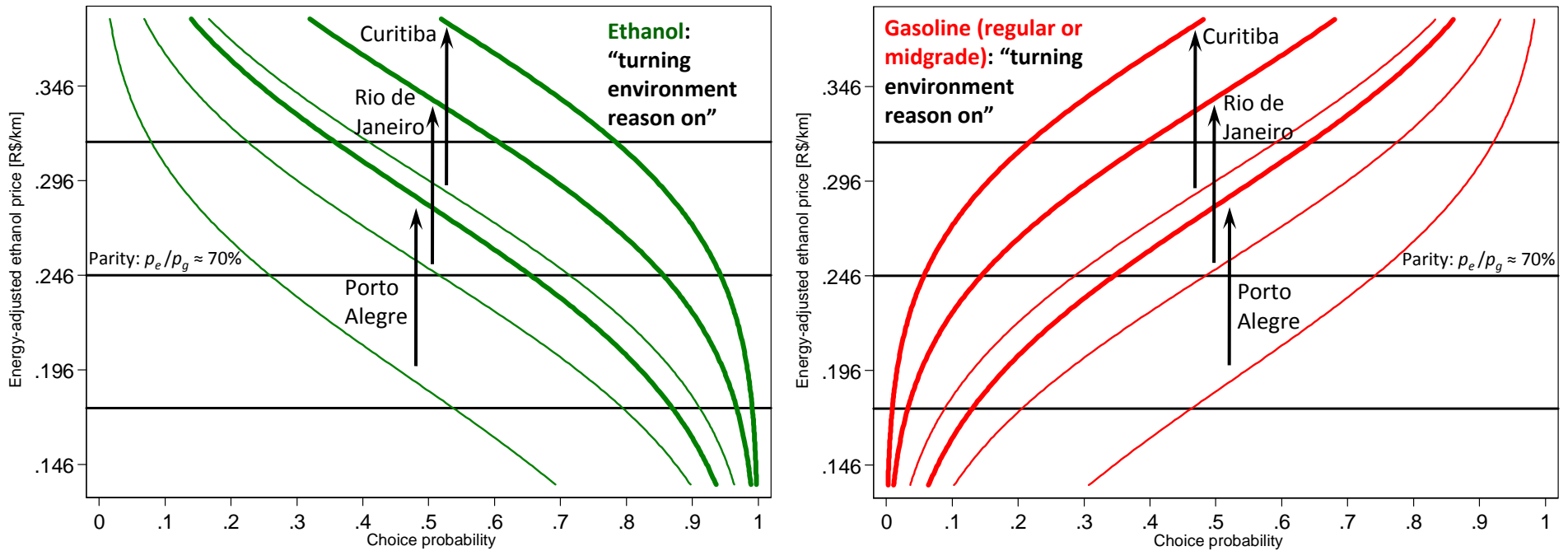


Figure 8: “Willingness to pay for greenness.” Choice probabilities for ethanol (left panel) and for gasoline (both varieties, right panel), when the environment-invoking main-reason dummy is switched from off (thin lines) to on (thick lines). Source: Specification II estimates. The plots consider a “median” motorist 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 motorist’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 kpl ratio of 70%).