Fueling Alternatives: Evidence from Real-World Driving Data

Ashley Langer* and Shaun McRae†

September 9, 2014

PRELIMINARY AND INCOMPLETE

Abstract

Development of a transportation system based on an alternative fuel requires both drivers to invest in vehicles and fueling stations to invest in infrastructure. We study the interaction between these decisions using real world driving data that identifies when and where drivers stop to purchase gasoline. We estimate a discrete choice model for the driver's choice of refueling location and show that drivers make a trade-off between the price of fuel and the time taken to deviate from their route. With these results, we simulate the willingness of drivers to adopt alternative fuel vehicles under different assumptions about the density of the alternative fueling network. The results suggest that the marginal cost of each alternative fuel and the fixed cost of alternative fuel vehicles and alternative fueling stations can dramatically change the market equilibria and alter the role of government in helping to create a self-sustaining alternative fuel market.

*Department of Economics, University of Arizona. Email: alanger@email.arizona.edu
†Department of Economics, University of Michigan. Email: sdmcrae@umich.edu.

1 Introduction

Gasoline and diesel remain the dominant fuels for transportation in the United States, with a combined 99.7 percent market share in 2011.[1] This is in spite of several decades of interest and investment in a succession of alternative fuels such as ethanol, methanol, hydrogen, natural gas, electricity, and propane. A fundamental problem faced by all alternatives to gasoline is the limited initial availability of refueling locations. Consumers are reluctant to purchase a vehicle that cannot be cheaply and conveniently refueled. Conversely, fueling station owners are reluctant to invest in expensive new infrastructure without an existing market of vehicle owners. This "chicken-and-egg" problem may suggest a role for government intervention if developing an alternative fuel market is considered socially desirable.

The central concept required for understanding this problem is the value to consumers of a more extensive fueling network. In this paper we develop a revealed preference approach to recover the distribution of consumer valuations for fueling station density. Our approach is based on observations of the willingness of drivers to undertake costly deviations from their optimal route in order to arrive at a cheaper gasoline station. This trade-off for drivers between lower fuel prices and greater time costs may vary based on the type of trip and the amount of fuel remaining in their tank. Based on the observed behavior of how real drivers shop for gasoline, we can simulate refueling behavior for drivers of alternative fuel vehicles facing a more limited set of refueling locations.

This paper is the first to use naturalistic driving data to analyze the refueling behavior of drivers. The data come from a year-long study conducted by the University of Michigan Transportation Research Institute (UMTRI) in which 108 drivers were provided experimental vehicles that included advanced crash-warning technology as well as monitoring equipment. During the experiment, detailed data from more than 220,000 miles of real-world driving were recorded. Drivers in the experiment were responsible for paying for their own gasoline during the 40 days that they drove the vehicle. By matching vehicle data to gasoline station locations, we identify nearly 800 refueling stops. Using daily station-level price data, we identify the price of gasoline at the station where the driver stopped as well as the price at nearby stations where the driver could have chosen to stop.

With this dataset we estimate a discrete choice model for the driver's choice of refueling location. We show that drivers make a trade-off between the price of fuel and the distance

[1]Source: Alternative Fuel Vehicle Data, Energy Information Administration. `http://www.eia.gov/renewable/afv/index.cfm`. This number includes ethanol and biodiesel used for blending with gasoline and diesel.

(or time) of the deviation from their route to arrive at the chosen gas station. People are willing to go further out of their way in order to pay less for gasoline. As in Houde (2012), these results give an estimated value of driver's time. However, unlike that paper, we have data on trips by individual drivers and so are able to demonstrate the heterogeneity in the estimated value of time across groups. We show, for example, that older drivers have a low implied value of time while drivers from high-income neighborhoods have the highest implied value of time. This heterogeneity has important implications for the design of alternative fuel policies. Drivers with characteristics that suggest a greater propensity to purchase an alternative fuel vehicle are also the drivers with a high value of time. This implies that they may be unwilling to deviate from their normal route in order to refuel their vehicle, even if the alternative fuel is significantly cheaper than gasoline.

We develop a conceptual model of an alternative fuel market in order to illustrate how drivers' willingness to drive out of their way plays a fundamental role in the evolution of the market. Based on the market share of alternative fuel vehicles, fueling stations invest in alternative fuel infrastructure until the zero profit condition is satisfied. Drivers choose between gasoline and alternative fuel vehicles in order to maximize utility, where the value of each type of vehicle depends on the fuel price, the vehicle price, and the fueling station density.[2] One stable equilibrium of this model is for there to be zero alternative fuel vehicles and zero alternative fuel stations. If other equilibria exist, there is generally a minimum threshold for either alternative fuel station or vehicle density, below which the market will collapse back to zero in the absence of government intervention. Such intervention may be necessary in order to "nudge" the market past this minimum threshold, but once the market is past this threshold it will converge to a higher equilibrium where government intervention is unnecessary.

We then use our discrete choice model of refueling demand to simulate the distribution of consumers' willingness-to-pay for an alternative fuel vehicle, for different levels of station density. This then allows us to calculate the alternative fuel vehicle penetration as a function of the alternative fuel station density, derived from the observed refueling choices of drivers. Combining this relationship with a simple model of firm entry, we demonstrate the potential market outcomes for two types of alternative fuel. The first type of alternative fuel ("ethanol") has a marginal cost that is only slightly lower than gasoline, and the capital costs of vehicles and infrastructure are very low. The second type of alternative fuel ("nat-

[2]We assume that all other vehicle characteristics are the same between gasoline and alternative fuel vehicles, although this assumption is straightforward to relax.

ural gas") has much lower marginal cost than gasoline, but the capital cost of vehicles and refueling infrastructure are high. We show that the market for ethanol requires fairly high station penetration before consumers are willing to adopt vehicles, but once the market is past this point it will continue to grow until all consumers prefer ethanol vehicles to gasoline vehicles. In contrast, for an alternative fuel with low marginal costs but high capital costs such as natural gas, some consumers will adopt vehicles at low station densities, even if those vehicles are more expensive than gasoline vehicles. If the natural gas market can become established, then it will converge to an equilibrium where a subset of gasoline stations also offer natural gas and some consumers choose to purchase the more expensive natural gas vehicles in order to gain access to the cheaper fuel, while other consumers prefer to purchase less expensive gasoline vehicles and pay more for fuel.

Our paper contributes to the large literature on gasoline demand. Given the amount of attention paid to gasoline prices, there is remarkably little empirical research on how consumers shop for gasoline. Many papers in this literature use aggregate data at a national or state level.[3] Apart from papers using aggregate data, other research estimates household-level demand for gasoline using data from household surveys.[4] All of these papers necessarily involve a high level of aggregation, either across consumers or, in the case of survey data, over time within a single household. This aggregation obscures the short-term dynamics in gasoline purchasing.[5] These short-term dynamics—where, when, and how much gasoline drivers buy—are of great theoretical and policy importance. To our knowledge, no existing research uses repeated transaction-level information for individual drivers. Without this data it is difficult to analyze patterns in gasoline shopping behavior such as repeat purchases at a single station or willingness to go out of the way for a lower price.

In addition, the policy application in this paper contributes to the literature on the adoption of alternative fuel vehicles. Empirical studies on the demand for ethanol include Anderson (2012) in a U.S. state and Salvo and Huse (2013) in Brazil. Corts (2010) shows

[3]For example, Hughes et al. (2006) use a monthly time-series of national gasoline shipments to estimate the price elasticity of gasoline demand for two periods, 1975–1980 and 2001–2006. They show that demand is significantly more inelastic in the more recent period.

[4]Wadud et al. (2010) use household data from the Consumer Expenditure Survey to estimate price elasticities for groups with different characteristics. Urban households with multiple vehicles are the most price-elastic, whereas rural households with a single vehicle are the least price-elastic. Puller and Greening (1999), also using Consumer Expenditure Survey data, decompose gasoline demand into demand for vehicle miles traveled and demand for fuel efficiency. More recently, Bento et al. (2009) use the National Household Travel Survey to jointly estimate demand for vehicle miles traveled and the discrete demand for vehicles, in a utility-theoretic framework.

[5]Levin et al. (2012) use daily city-level gasoline expenditure data to show that estimated elasticities are much higher with the daily data than using the same data aggregated to a monthly level.

how a government policy to increase the number of alternative fuel vehicles led to an increase in investment in alternative fuel infrastructure. Our analysis is more closely related to simulations that jointly model alternative fuel consumption, vehicle purchase, and infrastructure investment (Struben, 2006; Greaker and Heggedal, 2010; Chen, 2012). Other papers in this simulation-based literature model the optimal placement of alternative fuel infrastructure.

The remainder of the paper is organized as follows. Section 2 develops a simple theoretical model of investment in alternative fueling infrastructure and alternative fuel vehicles to highlight the important role played by the consumer valuation of the fueling station density. Section 3 section describes the driving and refueling data and provides a descriptive analysis of refueling behavior. Section 4 describes the empirical model for the driver's choice of gas station. Section 5 provides the results for this model. Section 6 then uses the results to simulate adoption of alternative fuel vehicles and investment in alternative fuel infrastructure for two different alternative fuels: ethanol and natural gas. Section 7 concludes.

2 Conceptual Framework

Many countries have enacted policies to encourage the adoption of alternative fuel vehicles. The stated reasons for these policies include reductions in greenhouse gas emissions, reductions in local air pollutants, and enhanced domestic energy security (Yeh, 2007). While all alternative fuel policies are generally "second-best" to taxing the externalities that arise from fuel combustion, we assume that the policy maker's objective is to create a domestic market for the alternative fuel at the minimum cost. The policy maker is faced with consumers who maximize utility and firms (fueling stations) who maximize profit. In order to focus on the interaction between consumer demand for alternative fuel vehicles and the provision of the alternative fuel at stations, we assume that the policy maker is choosing between subsidizing the fixed cost of alternative fuel vehicles and alternative fuel stations.[6]

The theoretical framework is developed from the model of Greaker and Heggedal (2010). There are S_G existing gasoline stations who each choose whether or not to enter the alternative fuel market. The number of alternative fuel stations that enter is $S_A = s_A S_G$, where s_A is the fraction of firms that install alternative fuel infrastructure. Entering the market involves paying a fixed cost F_A to cover the installation of new infrastructure to store and deliver the alternative fuel. Following the circular model of Salop (1979), we assume that

[6]Other policies to encourage the development of an alternative fuel market may include information campaigns, subsidies to research and design work of automakers, and subsidies to producers of alternative fuels. We do not consider these policies here.

stations are located equidistant from each other on a circle of circumference 1. Consumers with alternative fuel vehicles have a unit cost of travel of t. The cost of alternative fuel (expressed in dollars per refueling stop) is c_A.

The solution to this entry problem is standard (Salop, 1979). All alternative fuel stations charge the same price, p_A:

$$p_A = c_A + \frac{t}{S_A} \tag{1}$$

The price includes a markup over marginal cost but entry occurs until all alternative fuel stations earn zero profits:

$$S_A = \sqrt{\frac{V_A t}{r F_A}} \tag{2}$$

In this expression V_A is the number of alternative fuel vehicles and r is the cost of capital. The number of alternative fuel stations who enter is increasing in the number of alternative fuel vehicles and in the travel cost, and decreasing in the fixed cost of entry.

Consumers choose between two types of vehicles: alternative fuel vehicles A and gasoline vehicles G. If they choose a gasoline vehicle, then they can stop for fuel at any of the currently available stations. If they purchase an alternative fuel vehicle, then their station choice is limited to only those stations that have invested in the infrastructure to offer the alternative fuel.[7] The utility that consumer i receives from a vehicle of type j is described by equation (3).

$$U_{ij} = m_i \left(c - \frac{t}{4S_j} - p_j \right) - rK_j \tag{3}$$

In this expression m_i is the per period total distance traveled by consumer i scaled by the fuel tank capacity (that is, it is the number of refueling stops that the consumer makes per period). Driving is assumed to be fixed so that m_i does not change based on p_j. The constant c incorporates the value to the consumer of driving the distance m_i. As above, t is the cost of travel and S_j is the number of stations of type j. p_j is the cost of refueling with fuel j (per fueling stop) and K_j is the purchase price of a vehicle of type j.

Consumer i will choose an alternative fuel vehicle if $U_{iA} \geq U_{iG}$. Equation (4) provides

[7] Flex-fuel vehicles can run on either gasoline or the alternative fuel. Owners of these vehicles can stop at any of the stations and, for a subset of stations, have a choice between the two fuel types. At low levels of station penetration this greatly enhances the value of the vehicle to consumers. However, profits for firms are lower because increased competition from gasoline reduces markups. A future version of this paper will extend the analysis and simulation to incorporate flex-fuel vehicles.

the condition under which i will be indifferent between the two vehicles:

$$m_i \left(\frac{t}{4S_A} - p_A \right) - rK_A = m_i \left(\frac{t}{4S_G} - p_G \right) - rK_G \tag{4}$$

Substituting equation (1) into equation (4) and rearranging gives the distance travelled by a consumer who is indifferent between the vehicles, as in equation (5).

$$m^* = \frac{r(K_A - K_G)}{\frac{5t}{4}\left(\frac{1}{S_G} - \frac{1}{S_A}\right) + c_G - c_A} \tag{5}$$

Assume that alternative fuel vehicles are more expensive than gasoline vehicles ($K_A > K_G$) but the cost of the alternative fuel is less than the cost of gasoline ($c_G \geq c_A$). If m_i is uniformly distributed between $\underline{M}$ and $\overline{M}$ then the fraction of drivers who buy an alternative fuel vehicle, v_A, is given by equation (6).

$$v_A = \begin{cases} \frac{1}{\overline{M}-\underline{M}} \left[1 - \frac{r(K_A-K_G)}{\frac{5t}{4}\left(\frac{1}{S_G}-\frac{1}{S_A}\right)+c_G-c_A} \right] & \text{for } S_A > \frac{1}{\frac{1}{S_G}+\frac{4}{5t}(c_G-c_A)} \\ 0 \qquad \text{otherwise} \end{cases} \tag{6}$$

From equation (6) we see that increasing the number of alternative fuel stations S_A will increase the number of alternative fuel vehicles v_A. If the travel cost t is higher then increasing the size of the alternative fueling station network will have a larger effect on v_A. As expected, lower relative capital costs $K_A - K_G$ or lower relative fuel costs $c_A - c_G$ will also increase the number of alternative fuel vehicles. Finally, there is a threshold number of alternative fuel stations below which no one will buy an alternative fuel vehicle.

Figure 1 shows the two reaction functions from equations (2) and (6) in (v_A, s_A) space, based on one possible set of parameter values. Given the fraction of consumers who buy alternative fuel vehicles, v_A, equation (2) gives the number of firms that will enter. Given the fraction of stations with alternative fuel, s_A, equation (6) gives the number of alternative fuel vehicle purchases.

For the particular parameter values shown in Figure 1 there are three possible equilibria. Point A at (0,0) and Point C are stable equilibria. Point B is an unstable equilibrium. Any point slightly below or to the left of Point B will cause the market to collapse back to Point A. Any point slightly above or to the right of Point B will cause the market to grow until it reaches Point C. Therefore, shifting the market to just above Point B is required in order to create what Struben (2006) describes as a "self-sustaining equilibrium". At this point, without any subsidies, there are enough alternative fuel vehicles to justify investment in

alternative fueling stations and enough alternative fueling stations so that drivers are willing to purchase an alternative fuel vehicle.

The central challenge for alternative fuel policies is how to shift the market from the initial stable equilibrium A to point B. Subsidies for alternative fuel vehicle purchases will rotate the consumer line down while subsidies for alternative fueling stations will rotate the firm line up. Such policies will change the location of Point B but not, on their own, shift the market away from point A. In order to do this, the government may be required to "jump-start" the market by building stations or buying cars directly.

A key component for the analysis of alternative fuel policies is the value that the consumer places on a more extensive fueling station network, captured in this simple model by the variable t. This parameter enters both the firm and consumer reaction functions. It affects how far consumers are willing to go out of their way to buy a cheaper alternative to gasoline, which in turn determines the market power of fueling station owners. For a more realistic analysis we ideally want to capture the heterogeneity in where and how much consumers drive and how this affects their valuation of gasoline station locations. In the next section, we describe an unusual dataset of real-world driving behavior that can be used to better understand the trade-offs consumers make when choosing between gasoline stations.

3 Data

3.1 IVBSS experimental data

The driving data used in the paper are from the Integrated Vehicle-Based Safety Systems (IVBSS) study conducted by UMTRI from April 2009 to May 2010. During this study, identical vehicles were provided to 108 drivers in southeast Michigan for about seven weeks each. The objective of the study was to observe driver responses to modern safety equipment including lane-departure and collision warning systems. The drivers used the vehicles as if they were their own (including purchasing their own gasoline) and UMTRI collected a detailed dataset that included driving data such as location, speed, acceleration, heading, weather, and instantaneous fuel use, at a frequency of ten observations per second. Cameras in the vehicles captured video of the driver and the surrounding roadway, while radar identified nearby vehicles.

Table 1 provides characteristics of drivers in the sample.[8] Potential participants were

[8]The final sample for the analysis comprises 108 drivers. There were 117 drivers who were provided a vehicle. However, nine people were subsequently excluded from the sample due to non-compliance (such as

selected at random from all Michigan license holders living within a radius of approximately one hour's driving time from Ann Arbor. The sample was stratified to give equal numbers of males and females in three age categories: 20–30, 40–50, and 60–70. In order to be included in the experiment, participants were required to drive a minimum number of miles per day on average. As shown in the table, the mean distance for the experimental participants was 50.9 miles per day, equivalent to about 18,600 miles per year. In total, data from 6,275 hours (224,700 miles) of driving were observed and recorded.

Variables recorded by the monitoring equipment provide comprehensive, high-frequency information on vehicle operation and driver behavior. They include vehicle location, speed, heading, fuel consumption, and the distance to surrounding vehicles. One variable not recorded by the monitoring equipment was the fuel tank level. The amount of fuel remaining in the tank is the major factor that determines whether a driver stops for gasoline. We recovered an estimate of the fuel tank level using images from an in-car "over-the-shoulder" camera directed at the steering wheel and dashboard, combined with second-by-second fuel consumption data. The details of this procedure are described in the online data appendix.

3.2 Gas station stops and prices

We matched the vehicle locations from the driving data to a database of gasoline stations in order to identify potential refueling stops. The gas station data are from OPIS (Oil Price Information Service) and contain the name, brand, address, and approximate geographic coordinates for every gas station in Michigan and Ohio. For gas stations in southeast Michigan, we supplemented this information using aerial photography from Google Earth to add the exact latitude and longitude of the gas pumps and each of the station entrances. As shown in Figure 2, we identified every vehicle stop within a radius of 100 meters of a gas pump. We reviewed the left-side camera images for all of these potential stops. If the camera showed that the vehicle was stopped beside a gas pump (as in the figure), the stop was coded as a gasoline refueling stop.

Daily station-level prices for the entire sample period are from OPIS.[9] These provide the gas price paid by the driver in each of the stops.[10] The data also provide the gas price at

insufficient use of the vehicle or sharing the vehicle with another driver).

[9]The OPIS data only report the price for regular gasoline. The Honda Accords used for the experiment run on regular gasoline and we consider it unlikely that drivers used a different (and more expensive) gasoline grade.

[10]There are two potential issues with the OPIS price data. First, the data only report one price for each station and day. If the price changes during the day, then the reported price may not be the same as the price paid by the driver. Second, the data contain many missing daily price observations, particularly for

every alternative station where the driver could have chosen to stop instead.

Table 2 provides descriptive information for the 794 gas stops in Michigan and Ohio identified in the driving data.[11] The mean quantity of gasoline purchased at each stop is 8.3 gallons. The mean gas price paid by drivers in the sample is $2.59 per gallon, with a range from $1.97 to $2.96. Figure 3 shows the date and price of the gas stops, as well as the average daily gas price in southeast Michigan.

For each gas stop we calculated a measure of the excess time for the driver to arrive at the gas station. The calculation method for excess time is similar to that used by Houde (2012). Suppose the driver starts at location A, drives to the gas station at B, then continues on to location C. The excess time for the gas station stop is the fastest time for the route A to B to C, less the fastest time for the direct route from A to C. Travel times between points were calculated using Dijkstra's algorithm applied to street network data for Michigan and Ohio. Further details of the travel time calculation are described in the online data appendix.

Figure 4 shows the distribution of excess times to the gas stations that drivers stop at. Slightly more than half of gas stops have an excess time of 30 seconds or less. In most cases, these represent stops at gas stations along the route that the driver was on, with no deviation from the optimal route. As shown in Table 2, the median excess time for the chosen gas stations is 0.2 minutes.[12]

There are considerable differences across drivers in how they manage their fuel inventories. Figure 5 shows the distribution across drivers of their mean tank level after a gas stop.[13] About one third of drivers fill the tank at every gas stop, with a mean tank level of at least 90 percent after filling. The remaining drivers do not always fill the tank every time they buy gasoline. About 5 percent of drivers leave the gas station with their fuel tank less than half full.

As illustrated in Figure 5, we split drivers into terciles based on their mean tank level

gas stations in remote areas. We used several different interpolation mechanisms for the missing data. Final results are not sensitive to the interpolation method.

[11] A small number of gas stops were identified in states other than Michigan and Ohio. These stops are excluded from our analysis due to the lack of price data.

[12] Table 2 also shows summary statistics for the excess distances for the observed gasoline stops. Some of these excess distances are negative: the optimal distance from origin A to gas station B then to destination C is **less** than the optimal distance directly from A to C. This is a result of the optimal routes minimizing travel times, not travel distances. Traveling directly from A to C may be fastest on a highway but have a longer distance.

[13] In calculating the mean tank level after refueling, we exclude the final refueling stop for each driver, because this may not be representative of the driver's usual behavior. Although there was no requirement to do so, some drivers filled the tank immediately before returning the vehicle, while others returned the vehicle with the gas tank nearly empty.

after each gas stop. Table 4 shows summary statistics for refueling stops and drivers in each of the three groups. The right two columns show the difference in means between the highest and lowest terciles (roughly, the "fillers" and "non-fillers") and the p-value for a two-sided test that these are equal. Drivers who do not fill the tank are significantly more likely to be from the youngest age bracket and from lower-income neighborhoods. Only 14 percent of the "fillers" are aged between 20 and 30, compared to 47 percent of the "non-fillers". These demographic differences are consistent with a short-term liquidity constraint explanation for drivers not completely filling their tanks.

There are few differences between the three terciles in the timing and location of gas stops. Despite purchasing the smallest quantity, the average amount of time spent at the gas station is greatest for the drivers who do not fill their tank, although the difference is not statistically significant. The longer stops by this group may reflect greater use of cash payments inside the store instead of credit card payments at the pump.

Table 4 also shows that the "non-fillers" wait until the tank level is lower before they stop. The difference between the highest and lowest terciles of the mean tank level before stopping is 7 percent of the tank capacity. This is further illustrated by Figure 6, which shows for each tercile the observed probability of stopping for gasoline as a function of the fuel tank level. Very few drivers stop for gasoline when the fuel tank is half-full or more. The probability of stopping for gasoline rises steeply below one quarter of a tank, especially for the drivers who always fill their tank.

4 Empirical Approach

To understand how drivers would react to a limited choice of alternative fueling locations, we first investigate how drivers choose between gasoline fueling stations with different characteristics. In particular, we analyze how drivers make decisions about when and where to stop for gasoline based on the location of the station relative to the driver's route and the price of gasoline at different stations. Through this model we can better understand whether drivers are willing to drive out of their way to purchase less expensive gasoline and when drivers are most likely to consider stopping for gasoline. This then allows us to simulate how drivers would respond to having an alternative fuel available at a subset of stations, and what that behavior would imply for drivers' willingness-to-pay for alternative fuel vehicles relative to conventional gasoline vehicles.

We first present results for a full-information discrete choice model of the driver's choice

of gas station conditional on stopping on a particular trip. On a given trip $t = 1, ..., T$ each driver $i = 1, ...N$ has a choice of whether to stop at each of $k = 1, ..., K$ stations, with no outside option of not stopping. For the refueling trips only, each driver $i = 1, ...N$ has the choice of $k = 1, ..., K$ stations, and from each choice receives utility given by:

$$U_{ikt} = C + \alpha p_{kt} + X_{ikt}\beta + \varepsilon_{ikt} \tag{7}$$

where p_{kt} is the price at station k on the date of trip t, X_{ikt} is a vector of characteristics of the station including the time out of the way the driver would have to go to get to the stations, and ε_{ikt} is an extreme-value type 1 error.

We also estimate a nested logit model that incorporate the driver's decision to stop or not stop for gasoline on every trip and, conditional on stopping, the discrete choice of gas station. The utility that each driver receives for each choice $k = 0, ..., K$ is given by:

$$U_{ikt} = C + \alpha p_{kt} + X_{ikt}\beta + \xi_{i0t} + \varepsilon_{ikt} \tag{8}$$

ξ_{0t} is the value of not stopping on this trip, which we model as:

$$\xi_{i0t} = \gamma W_{i0t} \tag{9}$$

where W_{0t} is a vector of characteristics of the trip including the type of trip (commute or non-commute) and the amount of gasoline remaining in the fuel tank at the start of the trip. Not stopping is assumed to have a price of zero and a time driven out of the way of zero. Therefore, in this model, the driver simultaneously makes the choice of whether to stop on each trip and, if she stops, where to stop. The unobserved error term ε_{ikt} has a generalized extreme value distribution for which ε_{ikt} may be correlated with ε_{ik^*t} (k and k^* both greater than zero), while ε_{i0t} is uncorrelated with all other ε_{ikt}.

These models provide information on the trade-offs that drivers make when choosing between stations. Apart from the price of gasoline at each station, the most important characteristic for the driver is the location of the station relative to the driver's route. All else being equal, we would expect that stations located further from the driver's planned route are less likely to be chosen. For every trip and every potential gas station choice, the additional trip time in minutes that the driver would require to visit that gas station is calculated. Further details of this travel time calculation are described in the online data appendix. The trade-off between more expensive gasoline along the route or driving out of the way for cheaper gasoline provides an estimate of drivers' average value of time.

Another important geographical variable is the ease of access by drivers to the gas station. We expect that gas stations that require a driver to make a left-hand turn across oncoming traffic are less likely to be chosen than gas stations that the driver can enter by making a right-hand turn. The side of the road that the gas station is located on depends on the direction from which the driver approaches, and so will differ by trip. It is determined as part of the calculation of the excess time for the driver to arrive at the station. Other characteristics of gas stations in the analysis include the gasoline brand and whether the station is located near a highway exit.

Trip characteristics may also affect the decision to stop for gasoline and, conditional on stopping, the choice of gas station. One important trip characteristic is whether or not the driver is commuting to work. By coding work and home locations we were able to classify all trips into one of four "tour" types, where a tour is a chain of multiple trips that start and end at either the home or work location. Trips may be part of a tour that cycles from home-to-home, from home-to-work, from work-to-home, or from work-to-work. We classified as "commute" trips all trips that were part of a tour that ended at work: either home-to-work or work-to-work.

Drivers choose between a set of stations on each trip. The set of possible gas stations includes all gas stations within the same Core Based Statistical Area (CBSA) as the trip. Trips that cross several CBSAs are split into multiple trips, each within one CBSA. Most of the trips in the data take place in the Detroit-Warren-Dearborn CBSA. There are nearly 2,000 gas stations in this metropolitan area. For tractability, the choice set is defined based on a cutoff value for the excess time to the gas station on a given trip. In the main specification this cutoff is 10 minutes. As a robustness check the results for cutoffs of 5 minutes and 15 minutes are also presented.

We make several other important assumptions in estimating the model. First, we assume that if the driver had made a different choice about where to stop for gasoline, she would still have traveled to that station from the same starting location and would travel to the same destination after leaving the station. This assumption is required to calculate the excess time traveled to each station, including those the driver does not stop at. However, it rules out a scenario, for example, where a driver chooses to pick up coffee at a different coffee shop depending on which gas station she stops at.

In addition, we assume that the driver knows the location of all gasoline stations near her route, and, perhaps less plausibly, that she knows the current price at each of those stations when she starts her trip. In other work, we are further investigating the extent to which

drivers may not be aware of lower prices away from their normal routes and the value of this information. Finally, in this formulation consumers are unable to make dynamic decisions about where to stop for gas. The consumer makes a choice about whether and where to stop without any information about the characteristics of stations near the next trip.

5 Results

Table 5 shows the results for the two models described in Equations (7) and (8) above, without any heterogeneity across drivers. The first two columns show the results for the model conditional on stopping for gasoline, with and without an interaction between an indicator for trip type (commute or non-commute) with prices and excess time. The right two columns show the results for the nested logit model, in which drivers choose to stop or not to stop on each trip and, conditional on stopping, choose which gasoline station to stop at. For all specifications, both the price and the excess time coefficients have the expected sign and are statistically significant. That is, more expensive stations and stations further from the driver's route are less likely to be chosen.

Scaling the estimates by the price coefficient allows the magnitude of the effects to be more easily interpreted. Table 6 shows the results from scaling the excess time coefficients by the price coefficient to derive an implied value of drivers' time. For example, from the results in Column 1, each additional minute that the driver must deviate from their route to reach the chosen gas station is equivalent to 13.5 cents per gallon of gasoline (0.832/6.142). The average amount of gasoline purchased is 8 gallons, so the value of the time spent deviating from the route is \$1.08 per minute, or \$65 per hour.[14] This value is high relative to wages, but may reflect the fact that consumers are not actually aware of prices at stations that are not directly on their route, and this uncertainty means that they are not willing to drive out of their way even if the price they would pay is low. It is interesting to note that Houde (2012) also finds a high value of time of \$54 per hour using aggregate data on gasoline purchase (and the assumption that consumers perfectly know gasoline prices). Columns 2 and 4 from Table 6 show that the implied value of time for commute trips is more than double that for non-commute trips. Drivers on their way to work are much less likely to detour to a cheaper gas station.

For the nested logit model, the value of not stopping depends on the amount of fuel remaining in the tank and on the type of trip. Not stopping is highly preferred when the gas

[14] While this calculation does not currently take into account the cost of the fuel used to travel out of the way to the station, this cost would be very small relative to the high value of time.

tank is nearly full: the combination of the two coefficients is positive over the full range of tank levels observed in the data. Not stopping is also preferred for commute trips, although the estimate is not statistically significant. There may be other factors that affect the driver's decision to stop or not to stop, such as the expected distribution of prices that the driver will face on a future trip. These are not included in the current version of the model.

The other estimated coefficients are interesting. Gas stations with entrances that require only a right turn to enter are more likely to be chosen. The result in Column 1 suggest that a driver would pay an additional 6.3 cents/gallon to avoid crossing the road to a gas station directly opposite, equivalent to about $0.50 for the average purchase quantity. The gas brand coefficients (not reported) show that Costco, Speedway, and Meijer are the brands most likely to be chosen after controlling for price and location.[15] Sunoco, Valero, and 7-Eleven are the brands least likely to be chosen.

The disaggregate nature of our data allows us to further explore the heterogeneity in preferences for fueling stations across different types of drivers in our sample. Table 7 shows the results by gender and age group. Women appear to be more price and time sensitive than men, although only the time sensitivity difference is statistically significant. The breakdown by age group shows that sensitivity to price is much larger for 60–70 year-olds while sensitivity to time is smaller. These differences are both statistically significant. The implied value of time is different across age groups: $121/hour for the youngest group, $100/hour for the middle group, and $29/hour for the oldest group.

The choice set for all of the results presented so far includes all gas stations within the same CBSA as the trip that are no more than a 10-minute deviation from the driver's optimal route. We repeat the analysis for two alternative choice sets: all stations within a 5-minute deviation from the driver's route and all stations within a 15-minute deviation. These results (not shown) indicate that changing the cutoff value for excess time has no qualitative effect on the results.

6 Application to Alternative Fuel Vehicle Adoption

We use the results from the previous section on how and why drivers choose stations that are not on their shortest route to examine the difficulties faced by any alternative to gasoline. Producers of alternative fuel vehicles face a quintessential "chicken and the egg" problem:

[15]Some of these brands offer discount cards or promotions bundled with supermarket purchases. As a result, the price paid by some drivers may be less than the list price in our OPIS data. This may explain why these brands are preferred after controlling for (list) price and location.

no one will purchase an alternative fuel vehicle if they will not have anywhere to fill it and no one will convert part of their gas station to an alternative fuel if no consumers own alternative fuel vehicles. This has led many analysts to believe that government intervention to subsidize alternative fuel vehicles is necessary.

We use our model of gasoline station choice to calculate the value of networks of alternative fueling stations with different densities. In our model, drivers can choose to go further out of their way to buy cheaper gasoline. This is equivalent to a scenario in which an alternative fuel is cheaper than gasoline but the fueling infrastructure is not well developed.[16] For the time being, we will consider the trade off between the relative alternative fuel price and the density of alternative fuel stations, and construct willingness-to-pay differentials that allow us to quantify the cost to drivers of the limited availability of the alternative fuel.

6.1 Simulation

We use the estimates from the discrete choice model in Section 5 to simulate the effect of different levels of alternative fueling station penetration. In particular, we calculate the value to consumers of alternative fuel vehicles relative to gasoline vehicles if the alternative fuel is cheaper than gasoline but only available at a subset of gas stations. The drivers in our sample differ in the value they place on a more extensive fueling network, based on where and how much they drive. Therefore, we calculate the entire distribution of the differences in value between alternative fuel and gasoline vehicles over the drivers in our sample. We can then use this distribution to calculate, for a fixed price difference and station penetration, the proportion of drivers who would prefer an alternative fuel vehicle, for any given difference in vehicle purchase price.

We first calculate each driver's expected consumer surplus from the option of stopping at the complete set of gasoline stations on each trip over the entire period the driver has the vehicle. This consumer surplus is given by:

$$E[CS(gas)_i] = \frac{1}{\alpha} \sum_{t=1}^{T} \ln \sum_{j=0}^{J} V_{ijt} \tag{10}$$

$$\text{where} \qquad V_{ijt} = \alpha p_{jt} + X_{ijt}\beta + \xi_{i0t}$$

We can then compare this consumer surplus to the consumer surplus the driver would get

[16] Note the "cheap alternative fuel" could either mean that the alternative fuel itself is cheap or that the alternative fuel vehicle technology is sufficiently efficient to make the per-mile cost of driving lower than gasoline.

from driving an alternative fuel vehicle that can only stop for fuel at a subset of stations that have both gasoline and the alternative fuel. We label these stations $a = 1, ..., A$ and assume that the station has the same characteristics as it does in the real world where it only offers gasoline, with the exception that that alternative fuel may be offered at a different price than the gasoline. Using this notation, the consumer surplus a driver gets from purchasing alternative fuel over the period of study is given by:

$$E[CS(alt)_i] = \frac{1}{\alpha} \sum_{t=1}^{T} \ln \sum_{a=0}^{A} V_{iat} \tag{11}$$
$$\text{where} \qquad V_{iat} = \alpha p_{at} + X_{iat}\beta + \xi_{i0t}$$

To calculate this consumer surplus, we need to know the price that each station would charge for the alternative fuel given the set of competing stations offering the alternative fuel. To calculate the prices for each station, we use the simple model of firm markups as given in equation (1). In our data, there are, on average, 148 stations within 10 minutes of each route on which drivers stop for gas. Combining this with the assumed markup of 15 cents per gallon for gasoline and the fact that the average driver purchases 8 gallons of gas each time he stops gives us that the transportation cost, t, is 177.6. We then use this number to calculate the implied markups for each density of alternative fuel stations. For instance, if 10% of stations offer the alternative fuel, then, on the average trip where a driver stops he will pass within 10 minutes of 15 alternative fuel stations, and those stations will charge a markup of $\frac{t}{S_A} = \frac{177.6}{14.8} = \$12$ per stop or \$1.50 per gallon. We do this for each alternative fuel station density to find the markup curve for the alternative fuel, and then use the assumed difference in marginal cost to compute the equilibrium price difference between gasoline and the alternative fuel. Finally, we use this price difference to adjust the observed price at each station, which allows for unobserved differences in marginal costs across stations.[17]

Combining the consumer surplus from having the choice of gasoline stations with the consumer surplus from having the choice of alternative fuel stations gives us the driver's willingness-to-pay for a gasoline vehicle relative to an alternative fuel vehicle, given a set of stations which offer the alternative fuel:

$$E[WTP(alt)_i] = E[CS(gas)_i] - E[CS(alt)_i] \tag{12}$$

[17]For the time being we follow the simplified firm pricing model from section 2 to calculate alternative fuel prices. In continuing work we are planning to calculate station-specific markups directly from our demand model, which will allow markups to depend on the actual distance between stations offering the alternative fuel.

Because drivers in our sample only drove the vehicles for 40 days, we need to scale the willingness-to-pay up to the value consumers would have for a new vehicle purchase. We assume that the 40-day distribution of driving is representative of the driver's annual driving and multiply the willingness-to-pay from equation 12 by $365/40 = 9.125$ to get the annual willingness-to-pay. We then assume that both alternative fuel and gasoline vehicles survive for 14 years and that consumers have a discount rate of 8%. Both of these assumptions are consistent with the literature on new vehicle purchasing.

This approach clearly requires several substantial assumptions. First, we assume that the trips a driver would take in an alternative fuel vehicle would be identical to the trips she would take in a gasoline vehicle, except for a different choice of fueling station. Second, we assume that the experience of filling an alternative fuel vehicle is identical to the experience of filling a gasoline vehicle. For instance, this formulation rules out a situation where the driver must wait for an extended period of time for the vehicle to finish fueling. If that were the case, presumably the driver would make different decisions about when to refuel based on her ability to do other things like work, shop, or eat while the car is fueling. Third, in using the expectations operator, we are assuming that the driver is not aware of her iid random error draws ϵ_{ijt} when she makes her vehicle purchase decision. Fourth, we assume that markups for both gasoline and natural gas are constant across stations given the number of stations offering the fuel, and that the set of stations offering the alternative fuel are randomly allocated across all existing gasoline stations.

6.2 Market analysis

We combine the distributions of consumer willingness-to-pay for an alternative fuel vehicle with an extremely simple model of firm investment in alternative fueling stations. The profit of a firm that invests in an alternative fueling station is given by Equation (13):

$$\Pi = \frac{tqv_A V_G}{(s_A S_G)^2} - rF_A \tag{13}$$

Here t is the time cost of drivers, q is the annual purchase of alternative fuel per driver, v_A is the fraction of drivers who convert to alternative fuel, s_A is the fraction of stations that install alternative fueling equipment, F_A is the capital cost of the equipment, and r is the discount rate. The ratio V_G/S_G is the average number of drivers per fueling station.

Assuming that entry occurs until profits are zero, we can rearrange equation (13) to give the fraction of stations that install alternative fueling equipment as a function of the fraction

of drivers who convert to alternative fuel:

$$s_A = \frac{1}{S_G}\sqrt{\frac{tqv_A V_G}{rF_A}} \tag{14}$$

This equation is equivalent to Equation (2).

We choose realistic parameters for the analysis based on two different types of alternative fuel. The first is an alternative fuel with similar physical characteristics to gasoline, such as ethanol (E85). This fuel is assumed to be very similar to gasoline, but with marginal costs that are 10 cents per gallon equivalent cheaper than gasoline. Because of the similiarity in physical characteristics, capital costs to consumers or firms of converting to this alternative fuel are low. The vehicle price difference is assumed to be zero and the cost of installing the refueling equipment is assumed to be \$150,000 per station.[18]

For the second scenario, we consider an alternative fuel with very different physical characteristics from gasoline, such as natural gas. Because engines running on natural gas use the fuel more efficiently, the marginal cost differential for this fuel is assumed to be very large: \$1 per gallon equivalent. But because of the additional complexity in handling this fuel, the capital costs of the vehicles and the refueling equipment are assumed to be very large. The price of the vehicle is assumed to be \$2,000 higher than an otherwise identical gasoline vehicle and the cost of the refueling equipment is assumed to be \$2,000,000 per station.[19]

Using these assumptions and the estimated distributions of consumer willingness-to-pay for alternative fuel vehicles, we can construct figures similar to Figure 1 based on actual data. Figure 7 shows the result for ethanol. As before, (0,0) is a stable equilibrium. Starting from a very low number of ethanol cars and a low number of ethanol fueling stations, the market will eventually collapse to this point. Interestingly, it takes a fairly large number of ethanol stations in the market before any consumers are willing to purchase ethanol vehicles. This is because at lower station densities the market price of ethanol is actually above the price of gasoline: the markup charged by the few ethanol stations is actually larger than the 10 cen per gallon marginal cost difference. Once 60% of stations offer ethanol, competition

[18]We currently assume that once a consumer purchases a vehicle he can only use that type of fuel. This rules out flex-fuel vehicles that can run on either gasoline or ethanol. We are working on a scenario where ethanol and gasoline compete for flex-fuel vehicle owners' fuel purchases.

[19]For both alternative fuel scenarios, we choose parameters for the number of stations near each trip and the amount of fuel purchased from our data. The average number of stations within 10 minutes of a route in our data is 148, so we assume that $S_G = 148$. We use national data on the average number of vehicles per gasoline station in the US and set $V_G/S_G = 1637$. The average quantity of gasoline purchased in our data is 8 gallons, so $q = 8$.

in the ethanol market reduces the ethanol price below that of gasoline, and some consumers prefer ethanol vehicles to gasoline vehicles.

The firm and consumer reaction functions cross at an alternative fuel station penetration rate of just below 65% and a vehicle penetration rate of approximately 5%. This equilibrium occurs even at very low alternative fuel vehicle penetration rates because stations can cheaply enter the ethanol market to capture additional profits. In fact, once vehicle and station penetrations exceed 5% and 65% respectively, drivers are attracted by ethanol's lower prices and stations find it profitable to invest in ethanol fueling capability. Thus, even without any additional government subsidies, the ethanol market will continue to expand until it completely displaces gasoline. So although government intervention may be required to shift the market away from the initial equilibrium, there is a limit to the amount of subsidies required. When the ethanol market has expanded to either 65 percent of stations or 5 percent of vehicles, further growth will occur without any form of subsidy.

The result for natural gas, shown in Figure 8, looks very different. There is still a stable equilibrium at (0,0), but because natural gas has a substantially lower marginal cost than gasoline ($1 per gallon of gasoline equivalent), even at relatively low station densities natural gas is still priced lower than gasoline. When the share of stations offering natural gas exceeds 15%, some consumers start preferring alternative fuel vehicles to gasoline vehicles in order to take advantage of the fuel price savings. These are generally consumers who drive many miles in areas with a lot of stations. Their high mileage makes saving money on fuel attractive while their urban driving means that even with low alternative fuel station market share they are still close to an alternative fuel station when they need to purchase fuel.

The firm and consumer reaction functions for natural gas cross when the alternative fuel station penetration rate is just above 20% and the alternative fuel vehicle penetration rate is between 5 and 10% (point B). As with ethanol, this equilibrium is not stable, so if the alternative fuel station or vehicle penetration rate exceeds these levels, more consumers will want to purchase alternative fuel vehicles and more stations will want to offer the alternative fuel. Unlike with ethanol where the only stable equilibria are at (0,0) and (1,1), the natural gas simulation shows a second, stable equilibrium (point C) below full replacement of gasoline, where approximately 50% of stations offer natural gas and just over 40% of drivers purchase natural gas vehicles. Because of the extremely high fixed cost of offering natural gas, higher levels of natural gas vehicle adoption will not entice additional stations to enter the market. Similarly, because natural gas vehicles are more expensive than gasoline vehicles, there are some drivers who drive relatively little or pass relatively few stations who will always prefer

to purchase a gasoline vehicle rather than a natural gas vehicle. This means that natural gas and gasoline will share the transportation fuel market, with approximately half of gasoline stations offering natural gas and 40% of drivers purchasing more expensive vehicles that run on lower-cost but less-available natural gas.

Interestingly, the existence of these self-sustaining equilibria in both the natural gas and ethanol markets suggests that there is a role for government in encouraging the adoption of a socially-desirable alternative fuel. Because any level of station or vehicle penetration below point B in figure 8 will eventually collapse back to the equilibrium of (0,0), the investment that would be required by any one firm to build enough alternative fuel fueling stations to sustain the market would be enormous. For a fuel like ethanol, where the marginal cost differential is small, the number of alternative fuel stations the government would have to support before the market is self-sustaining is even higher. If the government decided that it wanted to intervene to create a stable alternative fuel market that did not require any government subsidies, it would need to assure that the market was above point B. To do this, the government might subsidize alternative fuel vehicles, which would shift the solid consumer function to the right, moving point B down and to the left. Additionally, if the government simultaneously subsidized stations' fixed cost of offering the alternative fuel, the dashed firm reaction function would shift up, moving point B farther down and to the left. Eventually, the government could cheaply purchase enough alternative fuel vehicles or fully finance enough alternative fuel stations to get the market above this lower, non-stable equilibrium. After this point, the subsidies could be completely removed and the market would expand on its own to the higher equilibrium at point C. A key outstanding question is the least-cost approach to pushing either the ethanol or natural gas market past the lower equilibrium.

To our knowledge, this is the first paper to use actual driver decisions about how to purchase fuel to understand how different alternative fuel markets could operate. While the existence of multiple equilibria in alternative fuel markets has been suggested theoretically and with simulations by Greaker and Heggedal (2010) and others, we are able to actually take the theoretical model to data to understand how drivers' willingness to trade off longer driving distances for lower gasoline prices and use that to show how realistic assumptions about the costs of natural gas could lead to multiple equilibria.

6.3 Sensitivity

Given the assumptions we have made in order to estimate these equilibria, it is important to understand how the location of the equilibria would change for a change in the basic assumptions. The two most important assumptions for each fuel are the cost difference between the alternative fuel vehicle and a conventional gasoline vehicle and the fixed cost for a station to add alternative fuel pumps. We show the sensitivity of the existence of an equilibrium at point B in Figure 1 by showing the approximate vehicle price difference and station fixed cost at which the B and C equilibria converge in Figures 9 - 12. These figures are all based off of Figure 1, with the blue dotted curve representing the stations' zero-profit frontier, with stations entering if we are to the right of the curve and the green solid curve representing the drivers' indifference between alternative fuel and gasoline vehicles, with drivers preferring alternative fuel vehicles to the left of the curve.

Figure 9 shows that if ethanol vehicles were $2,300 more expensive than gasoline vehicles (or their attributes were $2,300 less attractive to consumers than gasoline vehicles), the only equilibria would be at zero stations offering ethanol and 100% of stations offering ethanol. Interestingly, because of the low fixed cost to stations of offering ethanol, 100% of stations would offer ethanol even if only 15% of vehicles on the road were ethanol-only. Given that ethanol vehicles have characteristics very similar to gasoline vehicles, we think that a $2,300 price difference between the two is larger than we would expect, suggesting that there is an equilibrium beyond which the market is self-sustaining.

Figure 10 performs the same exercise, leaving the ethanol and gasoline vehicles at the same price, but raising the stations' fixed cost of offering ethanol until the non-zero ethanol market share equilibrium disappears. If we assumed that the fixed cost to a station of offering ethanol was $1.4 million, then there would be no equilibrium in the ethanol market other than at zero ethanol vehicle and zero ethanol station market share. Given that this is nearly 10 times higher than our estimate of the actual fixed cost of offering ethanol ($150,000), it seems unlikely that the real state of the world is one where there is no market equilibrium where at least some stations offer ethanol and some consumers prefer ethanol-only vehicles.

Similarly for natural gas, Figure 11 show that non-zero equilibria exist as long as the price (or value) difference between a natural gas vehicle and a gasoline vehicle is less than $8,000 if the fixed cost of adding natural gas fueling to an existing station is $1.5 million. Alternatively, if the vehicle price difference is $2,000, the fixed cost of offering natural gas at an existing station would need to be over $3 million for zero natural gas stations or vehicles to be the only equilibrium. As with ethanol, we believe that these assumptions are

fairly extreme, and therefore believe that under more realistic values there is an equilibrium beyond which the natural gas vehicle market is self-sustaining. Additionally, it is worth noting that the location of that equilibrium doesn't vary substantially in the Figures. In the base case (Figure 8), the minimum scale for a self-sustaining market was at approximately 20% of stations offering natural gas and 5% of vehicles running only on natural gas. Even in the extreme cases of $8,000 price premium for natural gas vehicles or $3 million fixed cost of offering natural gas fueling, the minimum scale for a self-sustaining equilibrium is well below 50% market penetration (approximately 35% of stations and 15% of vehicles if natural gas vehicles are much more expensive or 30% of stations and 20% of vehicles if the fixed cost of offering natural gas is much higher). This is somewhat different than the case of ethanol, where the minimum scale required for a self-sustaining ethanol market under extreme assumptions generally required a very high percentage of stations offering ethanol (over 80% in both cases and nearly 100% if ethanol vehicles are more expensive than gasoline vehicles) and a quite variable amount of vehicles running only on ethanol (15% of vehicles if ethanol vehicles are expensive relative to gasoline, but nearly 80% of vehicles if the fixed cost of adding ethanol to a station is high).

6.4 Policy Approaches to Creating a Self-Sustaining Alternative Fuel Market

Having shown that under reasonable assumptions there appears to be a (non-stable) equilibrium in both the ethanol vehicle and natural gas vehicle markets beyond which the market is self-sustaining, it is worth thinking about what the government might do if it wanted to help one of these markets reach this minimum self-sustaining scale. Without government intervention, no individual market participant unilaterally prefers to purchase an alternative fuel vehicle or invest in an alternative fueling station because there is not a sufficient network of vehicles or fueling stations. However, as shown earlier, if the government can subsidize alternative fuel vehicle purchasing and/or alternative fueling station investment to the point where the market is of a sufficient size, then the subsidies could be removed and the market would continue to function (and potentially expand) on its own.

The simplest way for the government to ensure that the minimum self-sustaining scale is achieved would be to either subsidize alternative fuel vehicles so that they are less expensive than gasoline vehicles or fully subsidize station investments in alternative fueling infrastructure.[20] In 2007, the most recent year for which both Census and National Highway

[20]We have checked our simulation, and either of these approaches would, indeed lead the market to reach

Administration data is available, there were approximately 120,000 gasoline stations and 250 million passenger vehicles in the U.S. If the government chose to use only vehicle subsidies, then achieving the equilibrium B from Figures 7 and 8 would cost approximately $1.25 billion for ethanol and $25 billion for natural gas. Alternatively, if the government chose to only subsidize alternative fuel stations, the achieving the same equilibria would cost $11.7 billion for ethanol and $36 billion for natural gas.[21] While these numbers seem large, it is important to remember that they are one-time investments, after which the market expands even in the absence of subsidies. This is in direct contrast to other energy subsidies, such as the current grants and tax credits that the government gives to renewable energy annually, which totaled $7.3 billion in 2013 alone, in addition to the $4.4 billion in tax credits for energy efficiency programs, and the $0.4 billion in tax credits for plug-in electric vehicles.[22]

There are pros and cons to subsidizing vehicles rather than stations. Clearly, the total cost of the subsidy is lower for vehicles than for stations. However, the minimum self-sustaining level of alternative fuel stations might be reached quickly since all stations could install alternative fueling capabilities at any time. On the other hand, less than 10% of passenger vehicles are new to the vehicle fleet in any given year. In order to achieve an alternative fuel vehicle market share of 5%, the government would either need to drastically increase the subsidy so that a much higher percentage of new vehicle purchasers choose alternative fuel vehicles or expect a very long time-frame before an alternative fuel market is self-sustaining.

7 Conclusion

Using a rich dataset of drivers' actual driving behavior, we have investigated the trade-offs that consumers make when deciding when and where to stop for gasoline. We find that consumers do trade-off the distance traveled out of the way for gasoline with the price, and that this relationship varies across drivers with different demographic characteristics.

We then use this model to better understand how consumers evaluate the choice of whether to purchase an alternative fuel vehicle given a limited number of stations selling the alternative fuel. We find that for an alternative fuel like ethanol with a small marginal cost

the minimum self-sustaining scale.

[21] These subsidy calculations are for a fixed subsidy to reach the equilibrium. The total cost would clearly be lower if the government instituted a dynamic subsidy scheme where the subsidy decreased in the total market share of alternative fuel vehicles or stations. Analyzing this dynamic subsidy path, particularly for fueling stations, is beyond the scope of this paper.

[22] Testimony of Terry Dinan, "Federal Financial Support for Fuels and Energy Technologies", `http://www.cbo.gov/sites/default/files/cbofiles/attachments/03-12-EnergyTechnologies.pdf`.

advantage over gasoline and low fixed costs of station entry, consumers will not choose to purchase alternative fuel vehicles until the alternative fuel station density is quite high, but that once a relatively small number of consumers have switched to the alternative fuel the market will likely move fully from gasoline to the alternative fuel. On the other hand, for an alternative fuel like natural gas with a large marginal cost advantage over gasoline and high fixed costs of station entry, consumers may purchase alternative fuel vehicles even at low station densities, but a stable equilibrium exists where both gasoline and the alternative fuel are sold.

This work suggests that there may be a role for government in helping a new alternative fuel to gain a foothold in the transportation fuel market, but that at least for some types of alternative fuels the government may be able to remove subsidies once the market reaches a sustainable level. We are continuing to work on quantifying the cost of government subsidies and purchasing to achieve this stable market equilibrium, which is necessary to understand whether the reduction in externalities from moving to the alternative fuel would be worth the cost of government intervention.

References

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

Bento, Antonio M, Lawrence H Goulder, Mark R Jacobsen, and Roger H von Haefen, "Distributional and Efficiency Impacts of Increased US Gasoline Taxes," *American Economic Review*, May 2009, *99* (3), 667–699.

Chen, Cuicui, "Working Towards a Future on Alternative Fuels: The Role of the Automotive Industry," 2012.

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

Greaker, Mads and Tom-Reiel Heggedal, "Lock-In and the Transition to Hydrogen Cars: Should Governments Intervene?," *The B.E. Journal of Economic Analysis & Policy*, 2010, *10* (1).

Houde, Jean-François, "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline," *American Economic Review*, August 2012, *102* (5), 2147–2182.

Hughes, Jonathan E, Christopher R Knittel, and Daniel Sperling, "Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand," *The Energy Journal*, 2006, *29* (1), 93–114.

Levin, Laurence, Matthew S. Lewis, and Frank A. Wolak, "High Frequency Evidence on the Demand for Gasoline," 2012.

Puller, Steven L and Lorna A Greening, "Household adjustment to gasoline price change: an analysis using 9 years of US survey data," *Energy Economics*, February 1999, *21* (1), 37–52.

Salop, Steven C, "Monopolistic competition with outside goods," *Bell Journal of Economics*, 1979, *10* (1), 141–156.

Salvo, Alberto and Cristian Huse, "Build it, but will they come? Evidence from consumer choice between gasoline and sugarcane ethanol," 2013.

Struben, Jero, "Identifying challenges for sustained adoption of alternative fuel vehicles and infrastructure," 2006.

Wadud, Zia, Daniel J. Graham, and Robert B. Noland, "Gasoline Demand with Heterogeneity in Household Responses," *The Energy Journal*, January 2010, *31* (1).

Yeh, Sonia, "An empirical analysis on the adoption of alternative fuel vehicles: The case of natural gas vehicles," *Energy Policy*, November 2007, *35* (11), 5865–5875.

Table 1: Descriptive statistics for drivers in IVBSS experiment

	N	Mean	Min	Median	Max
Number of trips	108	297	108	267	1,427
Days with vehicle	108	41	33	40	74
Total miles driven	108	2,080.2	591.3	1,921.2	5,854.6
Total hours of driving	108	58.1	21.9	53.0	131.5
Total gasoline used (gallons)	108	90.3	29.5	83.6	232.7
Miles per trip	108	7.5	2.6	7.1	14.3
Miles per day with vehicle	108	50.9	13.8	47.3	150.1
Miles per hour of driving	108	35.4	20.5	35.9	54.4
Miles per gallon	108	22.7	16.3	22.8	28.0
Number of gas stops	108	10	2	8	32
Gas used per stop (gallons)	108	9.1	3.6	9.3	16.4
Census block hh income (US$)	107	78,581	24,188	74,822	189,285
Demographic characteristics	Number	Proportion			
Male	54	0.50			
Female	54	0.50			
Age 20–30	36	0.33			
Age 40–50	36	0.33			
Age 60–70	36	0.33			
Commuter	68	0.63			
Non-commuter	40	0.37			

Table 2: Descriptive statistics for gasoline stops

	N	Mean	Min	Median	Max
Gas price ($/gallon)	794	2.59	1.97	2.60	2.96
Excess time from route (minutes)	794	1.1	0.0	0.2	13.2
Excess distance from route (miles)	794	0.4	-4.0	0.0	6.8
Amount purchased (gallons)	794	8.3	1.0	8.3	15.8
Purchase value ($)	794	21.5	2.7	21.4	43.8
Day and time of gas stop	Number	Proportion			
Day of stop = weekend	126	0.16			
12:00AM–6:00AM	22	0.03			
6:00AM–10:00AM	108	0.14			
10:00AM–2:00PM	189	0.24			
2:00PM–6:00PM	239	0.30			
6:00PM–12:00AM	236	0.30			
Gas station brand	Number	Proportion			
BP	146	0.18			
Speedway	108	0.14			
Marathon	103	0.13			
Mobil	75	0.09			
Sunoco	61	0.08			
Shell	43	0.05			
Meijer	50	0.06			
Other	208	0.26			
Gas station characteristics	Number	Proportion			
Right-turn entrance	466	0.59			
Right-turn exit	390	0.49			
Near highway exit	333	0.42			

Table 3: Descriptive statistics for trips

	N	Mean	Min	Median	Max
All trips					
Distance (miles)	31,001	7.2	0.0	3.4	166.5
Duration (minutes)	31,001	12.2	0.0	8.7	200.7
Gasoline used (gallons)	31,001	0.3	0.0	0.2	5.8
Miles per hour	30,766	29.4	0.1	25.9	120.0
Miles per gallon	30,807	18.8	0.0	19.4	49.9
Time between trips (minutes)	30,893	187.9	0.0	16.6	48378.6
Trips longer than 1 mile					
Distance (miles)	22,997	9.6	1.0	5.8	166.5
Duration (minutes)	22,997	15.6	0.2	12.4	200.7
Gasoline used (gallons)	22,997	0.4	0.0	0.3	5.8
Miles per hour	22,878	33.3	0.9	29.6	120.0
Miles per gallon	22,988	21.3	1.1	21.4	49.4
Time between trips (minutes)	22,893	213.1	0.0	26.1	48378.6

Table 4: Summary statistics for drivers and gas stops, by tercile of tank fill level

	Mean fill level (tercile)			Tercile 3 − Tercile 1	
	1	2	3	Value	p-value
Purchase characteristics					
Tank level after filling	0.58	0.81	0.93	0.34	0.000
Tank level before filling	0.27	0.30	0.34	0.07	0.000
Quantity purchased (L)	22.7	35.9	41.0	18.3	0.000
Purchase value ($)	13.57	20.32	24.10	10.53	0.000
Stop characteristics					
Gas stop length (min)	11.0	5.0	6.0	-5.0	0.305
Time out of route (min)	0.81	1.15	0.76	-0.05	0.717
Gas price ($/gallon)	2.30	2.17	2.23	-0.08	0.276
Price - average ($/gallon)	-0.10	-0.07	-0.04	0.07	0.045
Weekend (0/1)	0.26	0.27	0.27	0.01	0.811
Commute trip (0/1)	0.32	0.26	0.28	-0.04	0.272
Driver demographics					
Female (0/1)	0.53	0.50	0.47	-0.04	0.342
Age 20–30 (0/1)	0.47	0.39	0.14	-0.46	0.000
Age 40–50 (0/1)	0.28	0.33	0.39	0.13	0.001
Zip code income ($000)	73.2	76.3	86.4	15.6	0.000
Number of gas stops	382	284	254	.	.
Number of drivers	36	36	36	.	.

Note: Drivers in the sample are split into three groups based on their mean tank level after a gasoline stop. This table shows descriptive statistics for gas stops and drivers, for each of these groups. The final two columns show the difference in means between the first and third tercile of fill levels, and the p-value for a two-sided test that these means are equal.

Table 5: Gas Station Choice Results

	Conditional on stopping		Nested logit	
	(1)	(2)	(3)	(4)
Gas price ($/gallon)	-6.142*	-6.587*	-1.379*	-1.436*
	(0.656)	(0.727)	(0.199)	(0.217)
× Commute (0/1)		2.797		0.404
		(1.464)		(0.513)
Excess time (minutes)	-0.832*	-0.794*	-0.392*	-0.376*
	(0.045)	(0.047)	(0.030)	(0.029)
× Commute (0/1)		-0.322*		-0.163*
		(0.161)		(0.069)
Right-side arrive (0/1)	0.387*	0.387*	0.181*	0.182*
	(0.081)	(0.081)	(0.042)	(0.042)
Right-side leave (0/1)	0.031	0.026	0.026	0.024
	(0.080)	(0.080)	(0.039)	(0.039)
Highway exit (0/1)	-0.044	-0.047	0.017	0.016
	(0.104)	(0.104)	(0.042)	(0.042)
Brand effects	Y	Y	Y	Y
No stop equation				
Commute (0/1)			0.153	1.090
			(0.101)	(1.345)
Tank level (liters)			-0.019	-0.019
			(0.015)	(0.015)
Tank level squared			0.0027*	0.0027*
			(0.0003)	(0.0003)
Observations	85200	85200	1549938	1549938
No. of choice groups	794	794	24529	24529

Note: Each observation is a possible gas station that a driver could have chosen to stop at on a particular trip. Possible gas stations are within the same CBSA as the trip and no more than a 10-minute deviation from the driver's optimal route. For Columns 1 and 2, only trips for which the driver did stop at a gas station are included. Columns 3 and 4 include all trips and the option of not stopping. Gas brand dummies include the twelve largest brands and an "other" category for all other brands.

* $p < 0.05$ (two-tailed test for difference from zero).

Table 6: Implied Value of Time

	(1)	(2)	(3)	(4)
Value of time ($/hour)	65.04*		136.39*	
	(7.744)		(23.056)	
Value of non-commute time		57.86*		125.49*
		(7.166)		(21.931)
Value of commute time		141.40*		250.43*
		(52.509)		(120.352)
N	85200	85200	1549938	1549938

Note: Implied values of time are calculated from the price and excess time coefficients in the corresponding column of Table 5, based on an average purchase quantity per stop of 8 gallons.

* $p < 0.05$ (two-tailed test for difference from zero).

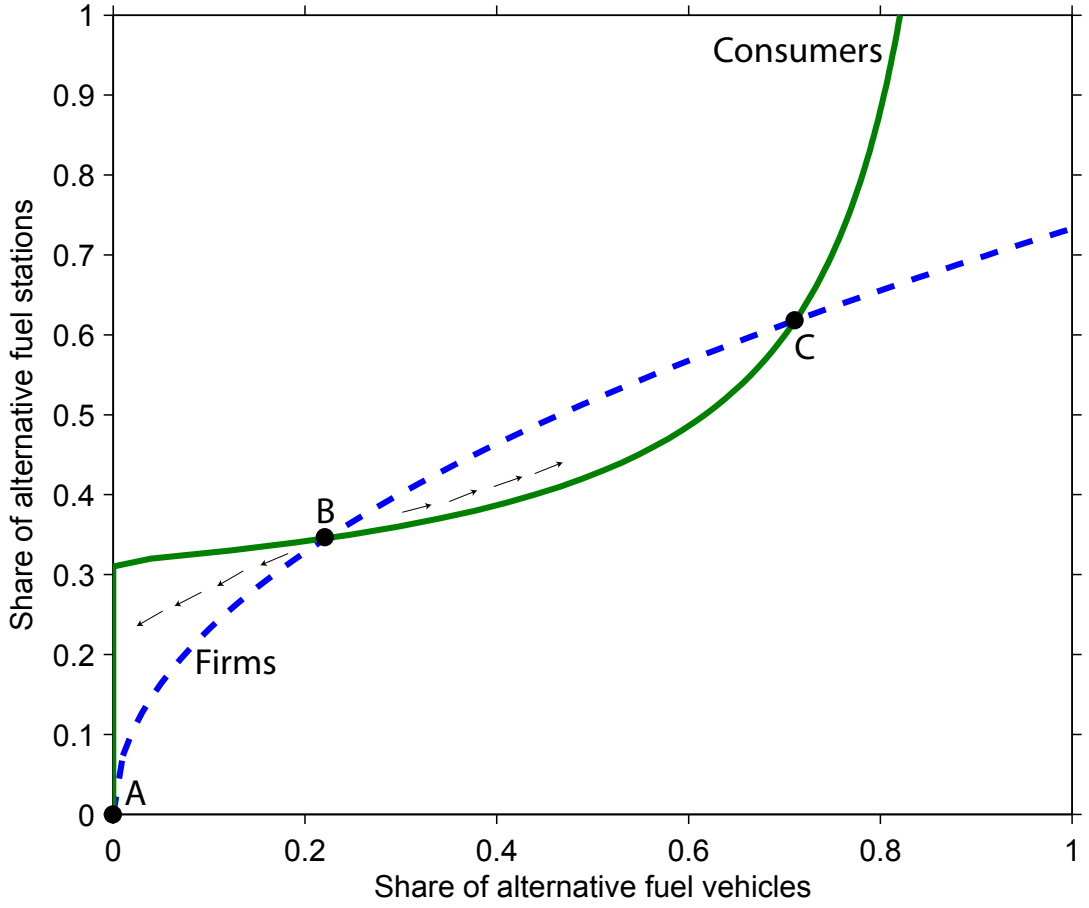
Table 7: Gas Station Choice Results by Demographic Group

	(1)	(2)	(3)
Gas price ($/gallon)	-6.142*	-6.087*	-3.381*
	(0.656)	(0.815)	(0.952)
× Female		-0.455	
		(1.275)	
× Age 40–50			-1.800
			(1.451)
× Age 60–70			-7.077*
			(1.498)
Excess time (minutes)	-0.832*	-0.719*	-0.852*
	(0.045)	(0.052)	(0.070)
× Female		-0.330*	
		(0.095)	
× Age 40–50			-0.228
			(0.122)
× Age 60–70			0.211*
			(0.095)
Right-side arrive (0/1)	0.387*	0.388*	0.356*
	(0.081)	(0.103)	(0.133)
× Female		0.016	
		(0.159)	
× Age 40–50			0.039
			(0.192)
× Age 60–70			0.107
			(0.190)
Right-side leave (0/1)	0.031	0.027	0.040
	(0.080)	(0.081)	(0.081)
Highway exit (0/1)	-0.044	-0.032	-0.076
	(0.104)	(0.104)	(0.104)
Brand effects	Y	Y	Y
Observations	85200	85200	85200
No. of choice groups	794	794	794

Note: As in Table 5, each observation is a possible gas station that a driver could have chosen to stop at on a particular trip. Possible gas stations are within the same CBSA as the trip and no more than a 10-minute deviation from the driver's optimal route. For all columns, only trips for which the driver did stop at a gas station are included.

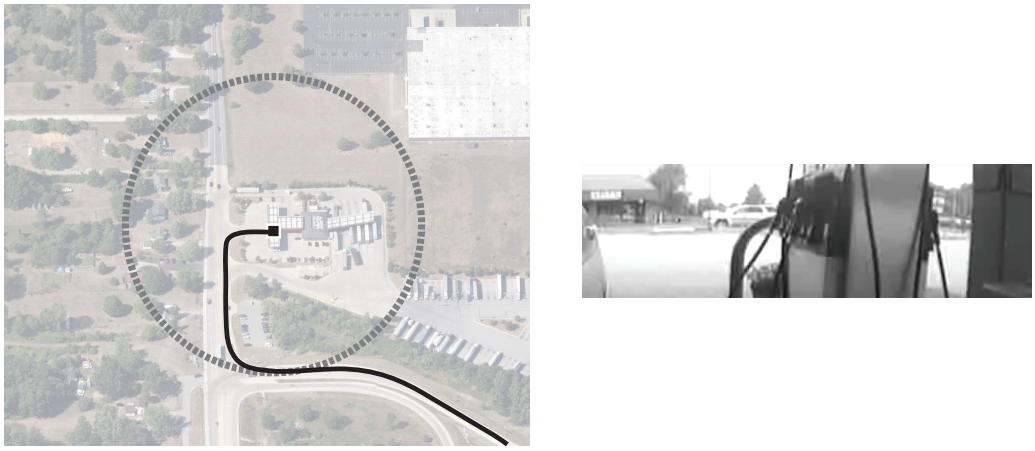
* $p < 0.05$ (two-tailed test for difference from zero).

Figure 1: Theoretical model for investment in alternative fuel vehicles and stations



The solid line in the figure is the share of consumers who will buy an alternative fuel vehicle as a function of the share of firms who invest in fueling infrastructure (from Equation (6)). The dashed line is the share of firms who invest as a function of the share of consumers with an alternative fuel vehicle (from Equation (2)). Parameter values are as follows: $F_A = 2{,}000{,}000$, annual $r = 0.08$, fuel price difference $c_G - c_A = 0.5$, number of gasoline stations $S_G = 55$, value of time $t = 50$, vehicle price difference $K_A - K_G = 2{,}000$, minimum stops per week $\underline{M} = 0.5$, maximum stops per week $\overline{M} = 2$, market size $V_G = 100{,}000$, purchase quantity per stop $= 8$.

Figure 2: Procedure for identifying gas station stops



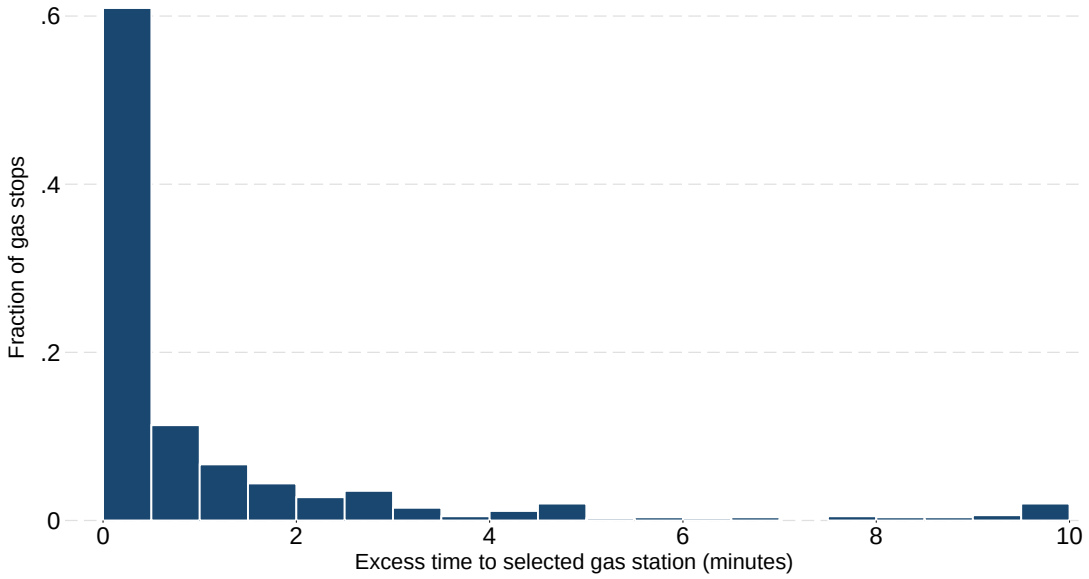
All vehicle stops within a 100-meter radius of gasoline station pumps were considered as possible refueling stops (left image). Images from the driver's side camera were used to confirm that the car was stopped at a gas pump (right image).

Figure 3: Average gasoline price and observed purchase price through sample period



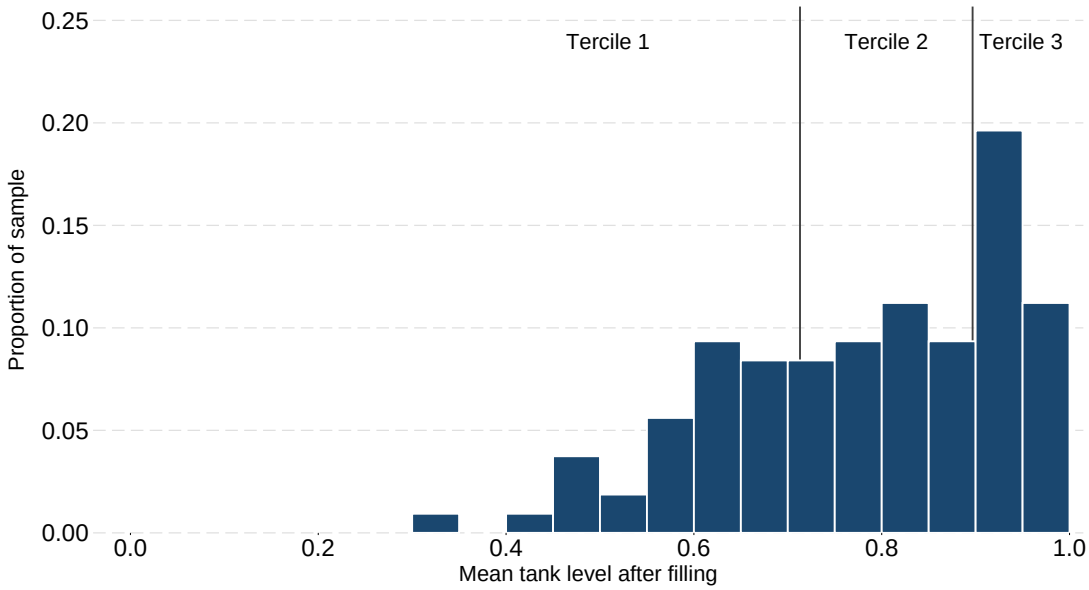
The thick black line shows the daily average gasoline price for the eight counties in southeast Michigan from the OPIS data. Each small dot represents the date and price of one of the gas station stops that we identify.

Figure 4: Distribution of excess time travelled to chosen gas station



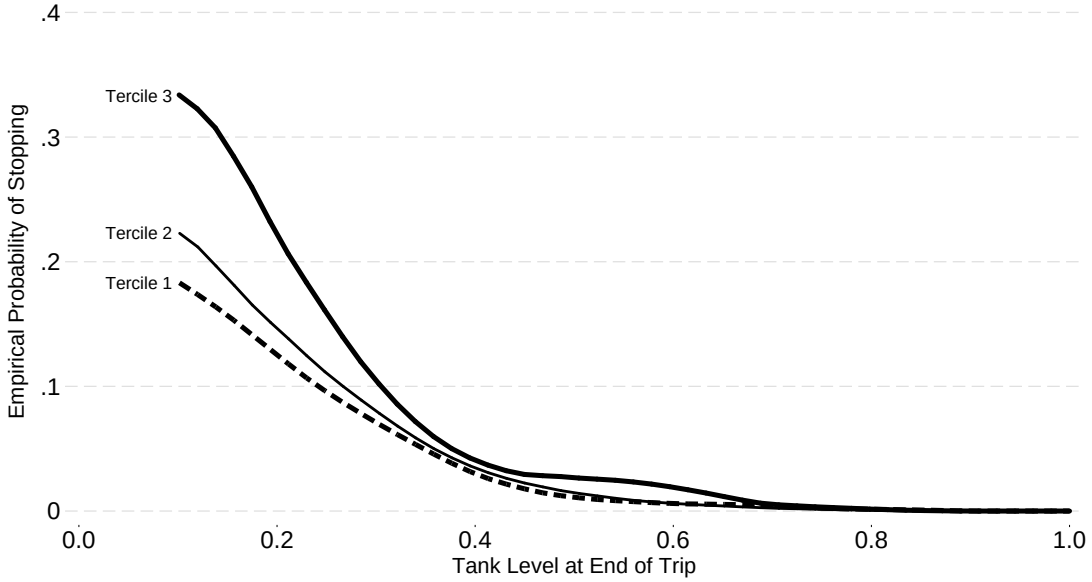
The map shows the distribution of excess times that drivers travelled away from their route in order to arrive at their chosen gas station. The calculation method for the excess times is described in the online data appendix. Times greater than 10 minutes are combined in the final right-side bar.

Figure 5: Distribution of tank levels after gasoline stops



The histogram shows the mean tank quantity, by driver, after a gasoline stop. The solid lines show the split of the driver sample into terciles, used for the summary statistics provided in Table **??**.

Figure 6: Probability of stopping for gas based on fuel tank level at end of trip



The graph shows local polynomial regressions of an indicator for stopping for gasoline at the end of a trip on the fuel tank level at the end of the trip (where 0 is empty and 1 is full). Each line shows the probability of stopping for different terciles of the mean fill quantity. Tercile 3 are the drivers who almost always fill their tank.

Figure 7: Market equilibrium for ethanol (small price difference, low capital cost)

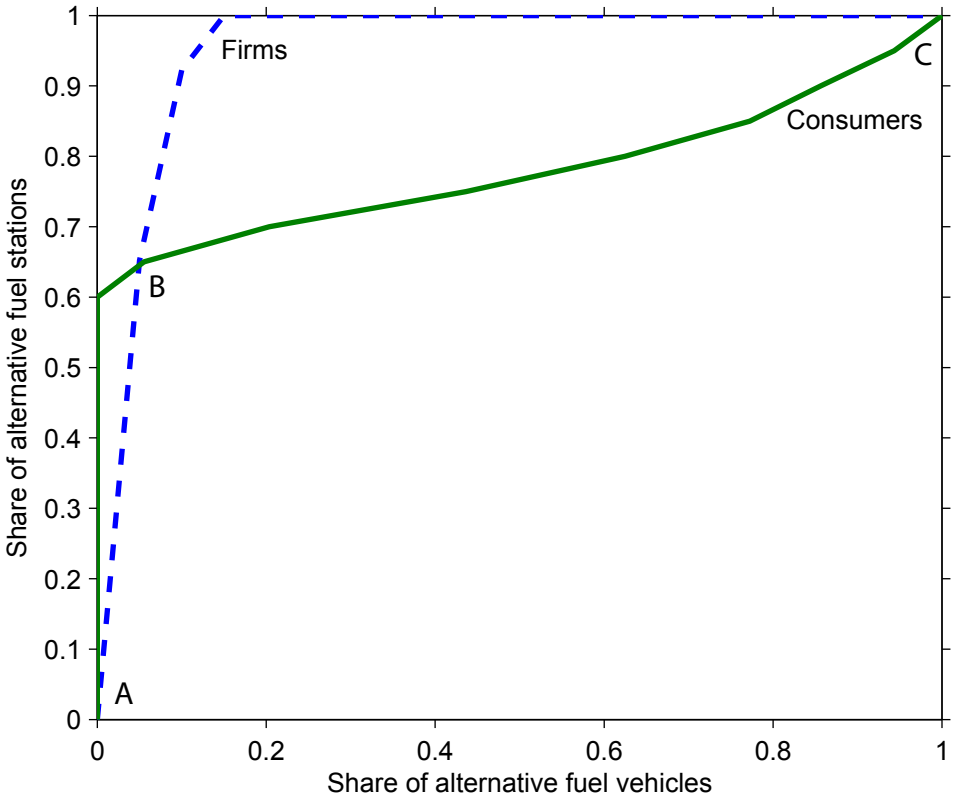


Figure 8: Market equilibrium for natural gas (large price difference, high capital cost)

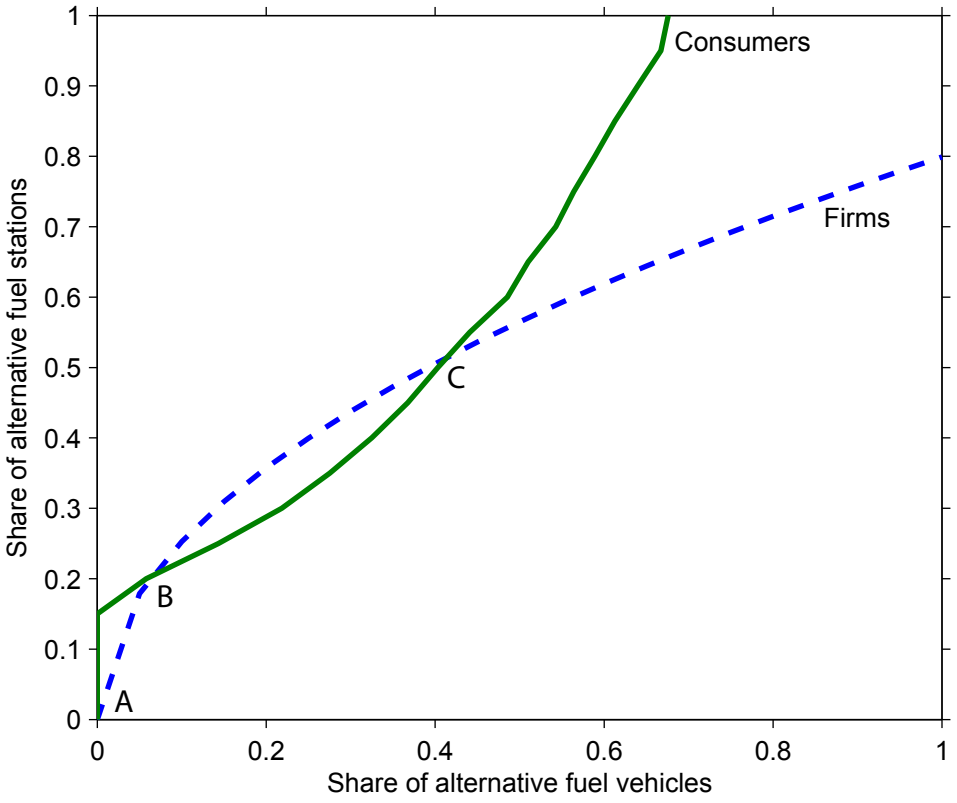


Figure 9: Ethanol sensitivity: $2,300 vehicle price difference

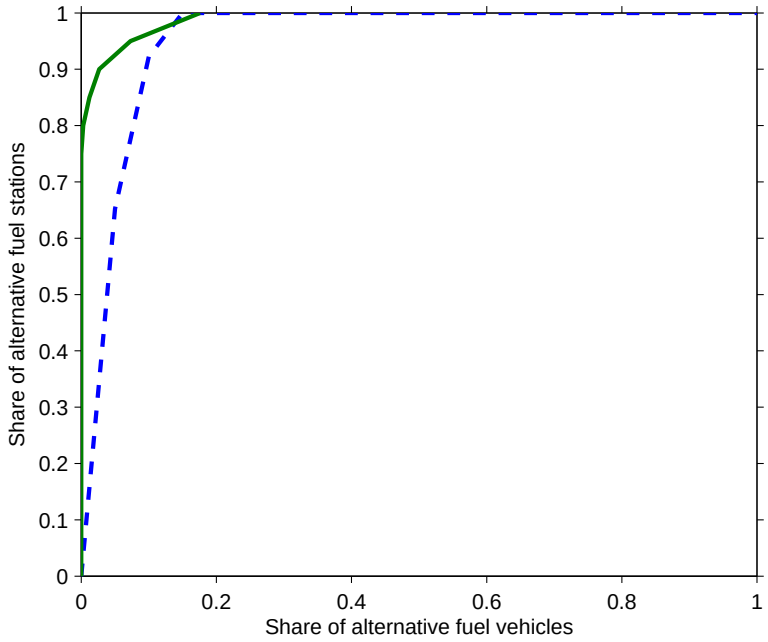


Figure 10: Ethanol sensitivity: $1.4 million fixed cost for stations

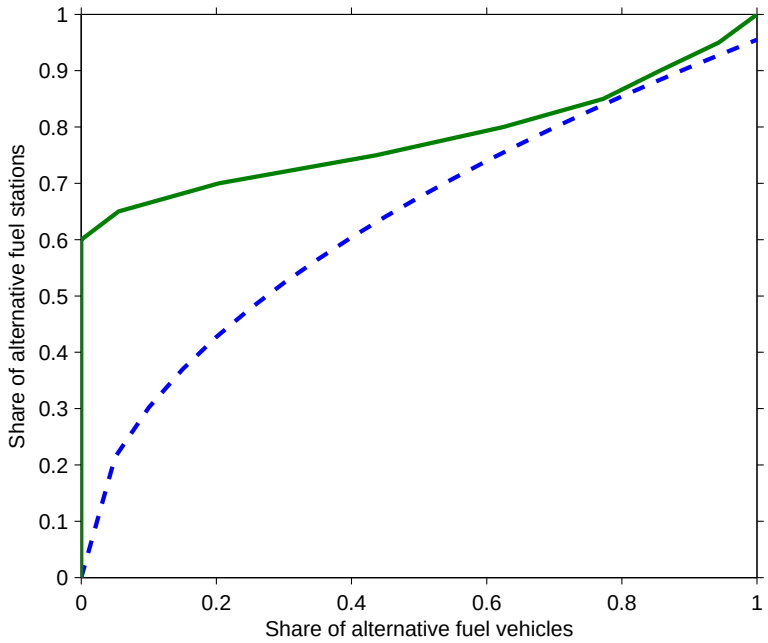


Figure 11: Natural gas sensitivity: $8,000 vehicle price difference

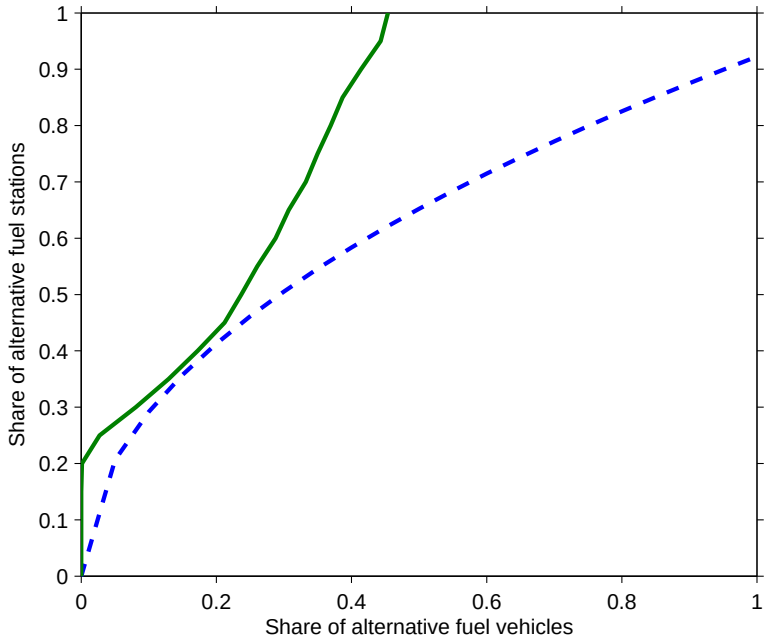


Figure 12: Natural gas sensitivity: $3 million fixed cost for stations

