Fueling Alternatives: Gas Station Choice and the Implications for Electric Charging*

Jackson Dorsey,[†] Ashley Langer,[‡] and Shaun McRae[§]

February 26, 2022

Abstract

This paper estimates an imperfect information discrete choice model of drivers' refueling preferences and analyzes the implications of these preferences for electric vehicle (EV) adoption. Drivers respond four times more to stations' long-run average prices than to current prices and value travel time at $27.54/hour. EV adopters with home charging receive $829 per vehicle in benefits from avoiding travel to gas stations, whereas refueling travel and waiting time costs increase by $9,169 for drivers without home charging. Increasing the charging speed of the existing network yields 4.7 times greater time savings than a proportional increase in the number of stations.

Keywords: gasoline purchasing, imperfect information, price perceptions, value of time, electric vehicles, charging stations

*We are grateful to MITRE, the U.S. Department of Energy, the Sloan Foundation, and the National Bureau of Economic Research for financial support. We also thank the University of Michigan Transportation Research Institute for data access. Finally, we thank Jean-Francois Houde, Dan Sacks, Shinsuke Tanaka, Matthijs Wildenbeest and seminar participants at University of Arizona, Indiana University, NBER, ASSA, UC Davis, University of Michigan, Northeast Workshop on Energy Policy and Environmental Economics for helpful comments.

[†]Indiana University, Kelley School of Business; jfdorsey@iu.edu

[‡]University of Arizona and NBER; alanger@email.arizona.edu

[§]ITAM; shaun.mcrae@itam.mx

Introduction

American households spend more each year on transportation than on healthcare, food, or entertainment—with fuel purchases alone accounting for over $2,000 of a typical household's annual expenditure (U.S. Bureau of Labor Statistics, 2020). On top of these private costs, gasoline combustion also creates harmful tailpipe emissions which has led to wide-ranging policy interventions including gasoline price constraints (Deacon and Sonstelie, 1985), fuel economy and emissions standards (Ito and Sallee, 2018), mandatory station price disclosure (Luco, 2019), and strategic petroleum reserves (Teisberg, 1981). More recently, policymakers have sought to encourage the transition from gasoline to electricity as the dominant transportation fuel. Because nearly all drivers currently refuel at gas stations, such an energy transition would require large-scale investment in electric vehicle (EV) charging stations. Recognizing this, the 2021 *Infrastructure Investment and Jobs Act* included funding to build half a million new charging stations.

Evaluating the effectiveness and efficiency of transportation fuel policies requires an understanding of how drivers make refueling decisions. However, we still know very little about these decisions or the preferences that underlie them. The lack of data on individual purchasing has meant that the academic literature has focused on understanding firm price setting (e.g. Borenstein et al., 1997; Byrne and De Roos, 2019; Luco, 2019) and using this pricing behavior to form inferences about drivers' behavior (e.g. Chandra and Tappata, 2011; Lewis, 2011; Yang and Ye, 2008), but has generally not been able to document drivers' trade-offs and preferences directly.[1] Yet these preferences are central to assessing a wide range of policies, from station price disclosure laws to subsidies for electric charging station investment.

In this paper we use uniquely detailed data on driving and refueling behavior to understand drivers' fueling preferences and the implications of these preferences for transportation electrification. We begin by using high-frequency GPS data on individuals' driving routes, gasoline refueling stops, and nearby station prices to document refueling patterns. Importantly, our data allow us to observe not only which gas stations drivers visit, but also the alternative gas stations they could have visited nearby their travel routes. We show that the median gas stop occurs within less than a one minute deviation from a driver's shortest route, despite the possibility of saving an average of $0.09 per gallon within a two minute deviation.

We further use these data to develop and estimate a discrete choice model of drivers' refueling decisions. Our model allows us to recover drivers' value of time, understand how they form perceptions of station prices, and simulate counterfactual choices with alternative fueling networks. On each trip, drivers in our model form expectations about the prices they would pay and the quantity of fuel they would purchase at each station, and then decide whether to stop for fuel and which station to stop at. The model captures fundamental

[1]Some papers have used information on commuting patterns and gas station prices (and sometimes quantities) to better understand drivers' preferences (e.g. Houde, 2012; Levin et al., 2017; Pennerstorfer et al., 2020), but these papers have generally not had information on individual drivers' trip characteristics or station choices.

features of drivers' refueling choice, including drivers' lack of perfect information about all stations, their increased incentive to stop for gas when their tank is low, their disincentive to travel further out of their way or pay higher fuel bills, and their preference for certain station brands. Our model incorporates imperfect information in two ways. First, because we observe where drivers have traveled in the past, we can limit drivers' choice sets to only those stations previously passed by the driver, in case drivers are unaware of other stations (Abaluck and Adams-Prassl, 2021; Goeree, 2008). Second, even if drivers are aware of stations, they may be imperfectly informed about station prices. We therefore explicitly allow drivers to form perceptions over the prices they would pay at each station that are a function of information such as the average price at a station and the current price at the station. Finally, to capture drivers' trade-off between total fuel expenditure and other station attributes, we model drivers as forming expectations over the quantity they would purchase at each station prior to making a station choice.

We identify drivers' value of travel time from their observed willingness to travel further from their routes in order to pay a lower expected gasoline price. We find that drivers have a high value of time of $27.54/hour, equivalent to 89% of the median wage of our sample group.[2] The value of time is a critical input into cost-benefit analyses of transportation policy (Small et al., 2005; Wolff, 2014), such as fuel economy standards, highway infrastructure investment, gasoline station zoning laws, or EV charging station subsidies. Notably, our value of time estimate is substantially higher than the estimate of 50% of the wage rate (White, 2016) used by the U.S Department of Transportation, which suggests that policymakers may be undervaluing the benefits of time-saving investments and regulations.

Our model also recovers an estimate of how consumers form perceptions of station prices, which may differ from actual prices due to imperfect information. We find that drivers respond four times more to each station's long-run average price than to its current price, which is consistent with consumers using average prices as a proxy for current prices. Yet perhaps surprisingly, we find that providing drivers with full information about prices would only increase consumer surplus by $0.02 per gallon or $0.15 per refueling stop. This low value of information results from the fact that drivers find it very costly to travel to cheaper stations that are away from their route Therefore, drivers gain little from learning about stations' prices located further from their routes. Overall, this set of findings supports recent research suggesting that mandatory price disclosure policies may result in limited benefits to consumers, particularly if the price disclosure facilitates firm collusion (Byrne and De Roos, 2019; Luco, 2019).

Beyond understanding drivers' preferences and the value of information in the gasoline market, our estimates help shed light on how refueling infrastructure affects drivers' incen-

[2] Median wage rates come from the Bureau of Labor Statistics' Occupational Employment and Wage Statistics for May 2010. We calculate the median wage as the average census-tract median wage where the drivers in our sample live.

tives to adopt EVs. Currently, homeowners are much more likely to purchase EVs than renters (Davis, 2019), likely because only 25% of apartment residents have access to EV charging at home (Ge et al., 2021; Traut et al., 2013). We find that the ability to charge overnight at home yields valuable time-savings to drivers. In particular, our estimates show that EV adopters with home charging access could save $829 in refueling time costs by avoiding trips to gas stations over the lifetime of a vehicle.

On the other hand, there are many drivers who would need to rely on public charging stations in order to charge their EVs. To understand drivers' value of public charging infrastructure, we merge the locations of all public charging stations with our geospatial data on drivers' trips. We see that, on average, drivers' routes in our data pass within five minutes of over 37 gasoline stations, but only 4.75 charging stations. Combining additional behavioral assumptions with our model's estimates, we can simulate where drivers would be likely to refuel an EV and the excess time associated with these stops. We find that in 2021, relying on EV charging stations instead of gasoline stations would increase excess time spent refueling from 2.5 minutes to 30.6 minutes per stop. More specifically, drivers spend an average of 1.8 minutes driving to gas stations and 0.7 minutes waiting at gas stations, whereas drivers spend an average of 30 minutes walking round-trip to EV charging stations and 0.6 minutes waiting at charging stations. Therefore, refueling an EV with public chargers would entail $9,169 in increased time costs over the life of a vehicle. This implies that, based on refueling time alone, drivers that can charge at home would value an EV at $9,998 ($9,169+$829) more than drivers without home charging. Thus, improving public charging infrastructure will be essential to encourage EV adoption by households without home charging.

Over the last decade, the number of U.S. charging stations has increased dramatically from around 1,000 in 2011 to nearly 50,000 in 2021, and the share of direct-current (DC) fast chargers has increased from 5% to 17% (DOE, 2022). We estimate that the increased station density and charging speed over this period reduced the excess refueling time for an EV from 121.6 minutes to 30.6 minutes per stop. Moreover, we calculate the elasticity of our excess refueling time measure with respect to the number of stations and the charging speed. This exercise reveals that increasing the speed of EV chargers would reduce drivers' refueling times by far more than increasing the number of charging stations, and that this gap has increased over time as more stations have entered the market. More specifically, a 1% increase in average charging speed in 2021 would generate 4.7 times more time-savings to drivers than a 1% increase in the number of charging stations. Although our exact refueling time estimates are sensitive to which assumptions we make, the headline comparison that investment in charging speed lowers drivers' refueling time more than investment in additional charging stations is robust across numerous alternative assumptions. This set of results indicates that policies that incentivize investment in DC fast charging infrastructure could yield greater benefits to drivers than investments in additional charging stations with slower charging technology (e.g. AC chargers).

Beyond the papers mentioned above, this work contributes to four different literatures. First, we extend the literature on how drivers purchase gasoline using data on actual refueling stops. One strand of this literature attempts to measure drivers' price elasticity of demand for gasoline in order to understand the effect of policies such as gasoline taxes (e.g. Hughes et al., 2006; Knittel and Tanaka, 2021; Levin et al., 2017). Other papers have investigated the role of search in gasoline purchasing and its effect on equilibrium market prices (e.g. Levin et al., 2019; Lewis and Marvel, 2011; Pennerstorfer et al., 2020). Finally, researchers have used the gasoline market to investigate other topics such as merger analysis (Houde, 2012), collusion (Byrne and De Roos, 2019; Lewis and Noel, 2011), and household budgeting (Hastings and Shapiro, 2013). This paper uses trip-level data to better understand when and where drivers choose to refuel and the information they use when making these choices, which has important implications for all three of these areas of the literature.

Second, this paper relates to work on how consumers' (potentially incorrect) perceptions of prices and other product attributes affect choices and welfare. Namely, in our model drivers make choices over whether and where to purchase gasoline without perfect information about current prices at every station they could choose. Earlier work in this literature showed how consumers' misperceptions of product characteristics can reduce their welfare (Allcott, 2013; Leggett, 2002; Liebman and Zeckhauser, 2004). Later work expanded this to include estimation of consumers' perceptions from their observed purchase decisions (Allcott and Knittel, 2019; Allcott and Taubinsky, 2015; Houde, 2018; Ito, 2014). In this paper, we estimate drivers' perceptions and calculate the value drivers would place on perfect information about station prices. This is important for understanding the potential value of mandatory price disclosure laws or websites and apps that provide widespread price information.

Third, we contribute to the established and growing literature on estimating consumers' value of time, particularly as it applies to transportation decisions. This literature has its roots in early empirical work such as Beesley (1965), but gained firm theoretical grounding with Oort (1969) which built on the broader work by Becker (1965). Work in this literature has recovered values of time from decisions over transportation modes (Lave, 1969), routes (Small et al., 2005), speeding behavior (Wolff, 2014), and rideshare choices (Buchholz et al., 2020; Goldszmidt et al., 2020). We provide what we believe is the first estimate of drivers' value of time from actual on-road refueling choices. Within this literature, the closest work to ours is Deacon and Sonstelie (1985) which uses a natural experiment that looks at drivers' willingness to wait in line at one gasoline station rather than pay a higher price without wait at another station. That paper also finds a high value of time (approximately equal to drivers' after-tax wages). Our value of time reinforces the other recent work in this literature (e.g. Goldszmidt et al., 2020) , which suggest that the Department of Transportation's current value of time—one-half of the wage rate—likely undervalues public investments and policies that provide time savings.

Finally, this paper adds to the important literature on the relationship between EV adoption and charging network investment. A series of theoretical and empirical papers have modeled the interaction between EV purchasing and charging station expansion, focusing on the “chicken and egg” problem of EV purchasers wanting assurances that they will be able to charge their vehicles, but charging stations being unwilling to enter until there is a critical mass of EVs on the road (e.g. Greaker and Heggedal, 2010; Li, 2017; Li, Jing, 2017; Springel, 2016). We build on this literature by providing the first empirically-based estimates of the refueling time costs associated with EV adoption and by measuring the elasticity of these time costs with respect to public investments in charging infrastructure.

We begin our analysis by describing our data and providing descriptive evidence on drivers' refueling choices in Section 2. We then discuss our empirical framework in Section 3 and present estimation results in Section 4. We apply our model to the valuation of EV charging, both for drivers who can charge at home and those who cannot in Section 5. Finally, Section 6 concludes.

2 Data

Our analysis relies on data from the University of Michigan's Transportation Research Institute (UMTRI) on individual driver behavior and fueling stops and data from the Oil Price Information Service (OPIS) on gasoline station locations and prices. In this section, we discuss each of these data sets in turn and provide detailed descriptive evidence on the characteristics of drivers' routes, station locations, and fueling choices. We also augment the UMTRI data with information on EV charging station locations from the Department of Energy's (DOE) Alternative Fuels Data Center to assess the value of EV charging infrastructure. We discuss the DOE's EV charging station data in Section 5.

2.1 IVBSS Experimental Data

We use driving data 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 approximately 40 days each.[3] 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, although the cars were given to the drivers with a full tank) and UMTRI collected a detailed dataset that included information such as GPS location, speed, acceleration, heading, weather, and fuel use, at a frequency of ten observations per second. Cameras in the vehicles captured video of the driver and the surrounding roadway.

[3]There were 117 drivers who were provided a vehicle. However, nine people were dropped from the sample due to non-compliance with the experimental guidelines. For example, drivers were disqualified for insufficient use of the vehicle or sharing the vehicle with another driver.

Appendix Table A.1 provides characteristics of drivers in the sample. Potential participants were recruited at random from all Michigan license holders living within a radius of approximately one hour's driving time from Ann Arbor with clean driving records. Of the drivers who expressed interest in the program, the final sample was stratified to give equal numbers of males and females in three age categories: 20–30, 40–50, and 60–70, who drove above a minimum number of miles per day on average.The table shows that the average driver lived in a census tract with median household income of approximately $64,000 per year which is above the median household income in Michigan, $54,379. However there was substantial variation across drivers, with census tract median household income ranging from below $20,000 to over $145,000.[4] Experimental participants drove 1,761 miles on average, equivalent to 51 miles per day (18,500 miles per year). In total, UMTRI collected data on 6,275 hours (224,700 miles) of driving.[5]

The data include comprehensive, high-frequency information on vehicle operation and driver behavior. In particular, we observe information about the time and location of every driving trip during the experiment. We define a unique driving trip as beginning each time a driver turns on their vehicle and ending when the driver shuts the vehicle down. The on-board computers document the starting and ending latitude and longitude associated with each trip. Additionally, we observe detailed data within each trip regarding the vehicle location, speed, heading, fuel consumption, and more, which we use at the one second level. These detailed data allow us to know exactly which route each driver took between the trip starting and ending locations. Appendix Table A.1 shows that the average driver made over 200 trips.

We aggregate the high frequency data to obtain a trip-level data set on fuel consumption and other variables of interest. Appendix Table A.2 provides details about the characteristics of trips in our sample. The median trip lasted 7.6 minutes (3.4 miles). Over half of trips lasted between 3 and 16 minutes, however some trips were substantially longer so the trip time distribution is right-skewed. We also see that the median trip begins only 2.4 miles from the driver's home. About 26% of trips take place during on weekends. 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 to refuel and how much gasoline they choose to purchase. 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. We describe details of this procedure in Appendix B. We estimate that, on average, drivers begin each trip with about 7 gallons (37%) remaining in the tank.

[4]We do not directly observe the drivers' home addresses. We infer drivers' "home address" as the most frequent end destination of that driver's trips. Appendix Figure A.1 shows the distribution of the drivers' median census tract income.

[5]Drivers were not compensated for their participation in the experiment other than through the use of the car and a nominal payment for completing baseline and endline surveys.

2.2 Refueling Choice, Gas Stations, and Prices

Although the IVBSS data provide extremely detailed information on the location of each trip, the on-board computers did not directly collect information about each driver's refueling choices. Therefore, we match the vehicle locations from the driving data to a database of gasoline stations in order to identify potential refueling stops. Gas station data come from the Oil Price Information Service (OPIS) and contain the name, brand, address, and approximate geographic coordinates for every gas station in Michigan and Ohio. 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 Appendix Figure A.2, we identified every vehicle stop within a radius of 100 meters of a gas pump, and then reviewed the rear-facing, 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. Drivers refueled an average of 8 times each during the experiment. Appendix Figure A.4 illustrates the locations of the 865 gas stops we identified in the data.

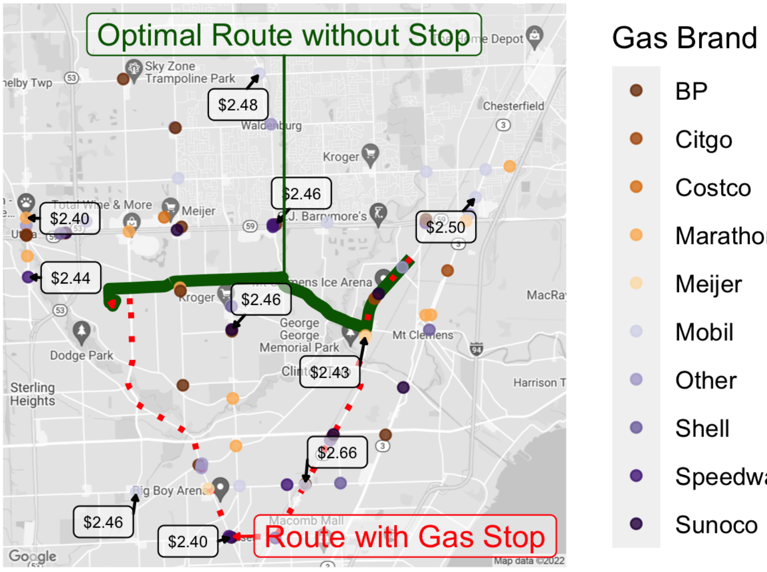
In addition to establishing gas station locations and brands, OPIS reports daily station-level prices for the entire sample period.[6] Thus, OPIS and vehicle GPS data together allow us to infer the gas prices paid by the driver at each of their stops.[7] Importantly, the OPIS data allows us to observe the gas price at every alternative station where the driver could have chosen to stop instead. Appendix Figure A.3 shows the date and price of the gas stops, as well as the average daily gas price across all stations in Michigan and Ohio.

For each driving trip we also calculate the excess time required for the driver to travel from the trip starting location to each potential gas station and then to the trip ending location, relative to the most direct route between the trip starting and ending locations. The calculation method for excess time is similar to that used by Houde (2012). Suppose the trip originates at location A, proceeds to a gas station at location B, then continues on to location C. The excess time for the gas station stop at location B 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 the Open Source Routing Machine (OSRM) applied to Open Street Map data for Michigan and Ohio (Luxen and Vetter, 2011). Figure 1 illustrates our excess-time (and excess distance) calculation for an example trip in Sterling Heights, MI. The green line plots the fastest route from the driver's origin to their destination without a gas stop. The dotted red line shows the quickest route if the driver were to refuel at a Sunoco station located to the southwest of their destination. Therefore, the excess time associated with choosing to stop

[6] 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 given that they do not actually own the vehicles.

[7] 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 some missing daily price observations, particularly for gas stations in remote areas. We used several different interpolation mechanisms for the missing data. Final results are not sensitive to the interpolation method.

Figure 1: Example Trip Route including Nearby Refueling Options



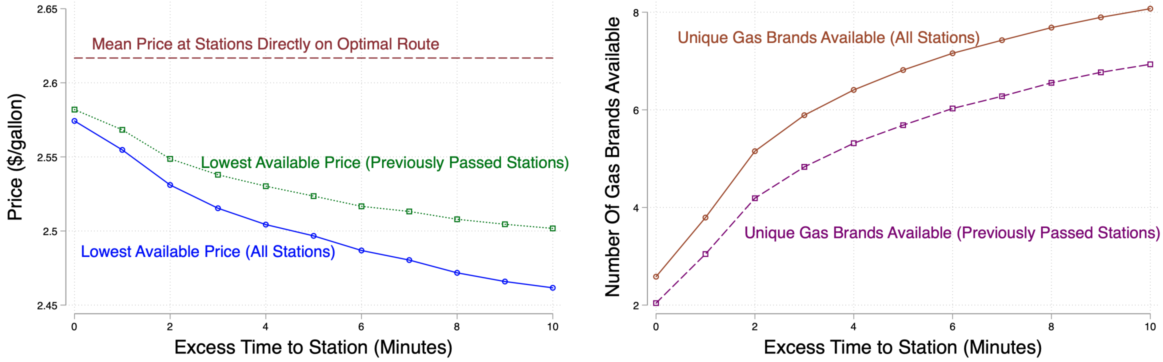
Notes: The green line shows the optimal direct route from origin to destination for a trip taken by driver #58 in Sterling Heights, MI. The red line shows the implied route if the driver chose to stop at a Sunoco station located to the southeast of the final destination. Each point shows the location of gas stations near the driver's route. Prices are labeled for a random subset of the stations for clarity.

at this Sunoco station would be the time to travel the red route minus the time to travel the green route. We perform this calculation separately for each station to determine the excess time needed to visit each possible gas station within 20 minutes of the route.

The bottom of Appendix Table A.2 provides a summary of the number of stations available to drivers during driving trips. Drivers typically have numerous refueling options. On an average trip, a driver could access six gas stations within just one minute of their route, nearly 25 stations within 5 minutes, and over 100 stations within a 20-minute diversion. This number of stations seems reasonable given the suburban Detroit region where most of our data's driving occurs.

Figure 2 demonstrates that driving farther from their route provides drivers improved options in terms of gasoline prices and brand options. In the left-side panel, the red dashed line shows the mean price a driver would pay at stations located within 15 seconds of their route. The blue line shows how the lowest price available (averaged over trips) changes with each incremental minute further away from the optimal route. A driver can find lower prices with each incremental minute of excess time for two reasons: 1) excess time from route is likely to be correlated with overall traffic flows and stations located on more common routes can charge higher prices, and 2) the driver can access a larger set of stations with longer diversions so the expected minimum price available would fall even if price and excess time were not correlated. We see that the largest incremental saving opportunities occur within

Figure 2: Price Frontier and Brand Variety by Excess Time From Drivers' Optimal Route



Notes: In the left panel, each point is the lowest available price offered (mean) by stations given the excess time it would take the driver to reach those station on a trip where the driver stops to purchase gas. The dashed line shows the mean price that a driver would pay among stations located within 15 seconds of their optimal route. The right panel plots the mean number of unique gas brands that are available to drivers given the excess time it would take the driver to reach those stations.

the first couple of minutes deviation from the route: a driver would save about $0.06/gallon at the lowest-priced station one minute away from their route relative to the average price available directly along their route. However, possible savings from driving further off route occur at a decreasing rate. In particular, a driver needs to travel an additional three minutes to save an additional $0.06/gallon (four total minutes to save $0.12/gallon total). The right panel of Figure 2 shows that additional driving provides access to more brand options in addition to lower prices. An average driver can reach three additional gas station brands by driving two minutes off their route and can access six additional brands by driving 10 minutes off their route.

During the period of our analysis, sites like *gasbuddy.com* were in their infancy and were not widely known or used, so information about gas stations and prices would likely need to be accumulated by actually passing the stations, or by discussing gas prices with others. As a consequence, drivers may not always be aware of all stations and current prices near their route. Using the high-frequency nature of the IVBSS data, we identify gas stations that are more likely to be known to drivers: the set of gas stations that drivers have previously passed in our data. We identify a station as being previously passed if the driver has passed within one-tenth of a mile of the station location during any previous trip. The green line in the left panel of Figure 2 illustrates that conditional on excess time, drivers are likely to pay a somewhat higher price for gas if they only consider the set of stations that they have previously passed. The right panel shows that drivers will also choose from a smaller set of gas brands if they only consider stations that they have previously passed.

Table 1 provides descriptive information for the 865 gas stops in Michigan and Ohio that we observe in the driving data.[8] Most drivers stop when their tanks are close to empty, with

[8] 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.

75% of stops occurring when the vehicle's tank level is below four gallons, and the mean quantity of gasoline purchased at each stop is slightly under eight gallons. The mean gas price paid by drivers in the sample is \$2.60 per gallon, with an interquartile range of \$2.50 to \$2.70. Drivers are equally likely to refuel during weekends relative to weekdays (weekends represent 26% of stops and 26% of trips).

Table 1: Summary of Refueling Stops

	Mean	SD	Pct25	Median	Pct75
Tank Level at Start of Trip (gallons)	2.70	2.23	1.07	2.24	3.99
Purchase Quantity (gallons)	7.82	4.89	4.10	7.72	11.64
Price Paid (\$/gal.)	2.60	0.16	2.50	2.60	2.70
Weekend (0,1)	0.26	0.44	0.00	0.00	1.00
Station to Home $\leq$ 10 Miles (0,1)	0.68	0.47	0.00	1.00	1.00
Passed Station Previously (0,1)	0.86	0.35	1.00	1.00	1.00
Refueled at Station Previously (0,1)	0.18	0.38	0.00	0.00	0.00
Excess Time (min)	1.88	2.97	0.02	0.67	2.45
Excess Distance (miles)	0.83	1.89	0.00	0.12	0.86
Observations	865				

Notes: Summary statistics are reported across all refueling stops in Michigan and Ohio made by all drivers during the experiment. Pct25 and Pct75 are the 25th and 75th percentiles, respectively.

Beyond providing information about the trips on which drivers stop for gas, Table 1 describes the stations at which drivers choose to refill. Drivers are more likely to stop at stations that they are likely to be familiar with: drivers pick stations located within 10 miles of their home 68% of the time and choose stations that they have passed previously passed during our sample period 86% of the time. Moreover, drivers return to the same station as they refueled most recently 18% of the time.[9]

Beyond stopping at stations they are familiar with, drivers tend to refuel at stations that are relatively close to their routes: the mean excess time for selected stations was 1.9 minutes (0.83 miles). Figure 3a shows the distribution of excess times to the gas stations that drivers choose to stop at (in green), compared to the unconditional distribution of excess times to all potential gas stations (in gray).[10] Gas stations are roughly uniformly distributed across excess times, but unsurprisingly, drivers are much more likely to stop at stations close to their routes. Indeed, nearly half of gas stops occur within 1 minute or less of the driver's route. Moreover, 40% of stops occur directly along the driver's route (within less than a 15 second deviation from the optimal route). Although the typical gas stop occurs relatively close to a driver's route, Figure 3b shows that drivers usually *do not* pick the most convenient station available. In particular, drivers almost always drive further out of their way for gas than they need to.

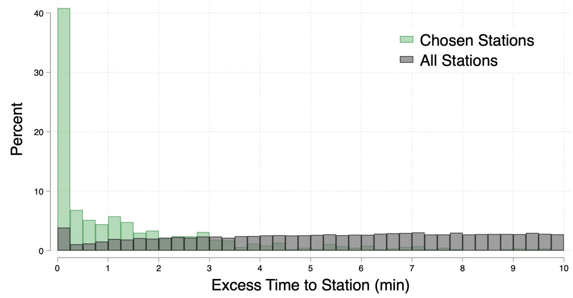
[9]BP and Speedway were the most common brands choices with an 20% and 15% share of the gas purchases in our data, respectively. Appendix Table A.3 provides the share of stops at each brand in our data. Smaller regional brands and unbranded stations were chosen at 13% of stops.

[10]We exclude stations more than 10 minutes away from the optimal route in this graphic for clarity.

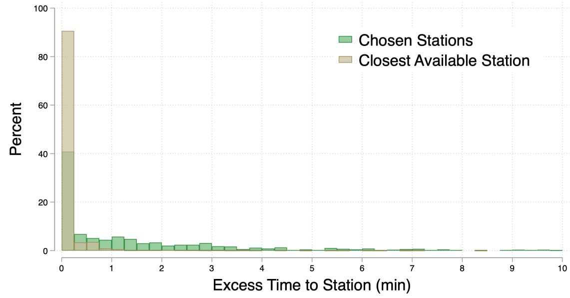
Figure 3b shows that drivers could have stopped at a station within 15 seconds of their route on over 90% of trips but they choose the most conveniently located stations only about 40% of the time. This pattern of drivers choosing stations that are slightly out of the way but not too far off their route is consistent with the decreasing marginal savings from driving further that we saw in Figure 2.

Drivers' final choice of where to refuel will depend on which stations the driver is aware of and actively considering. A driver previously passing a gas station is a useful indicator that a driver is aware of that station. Table 1 reveals that nearly 90% of refueling stops occur at stations that drivers passed previously. Figure 3c plots the distribution of excess time to stations that drivers previously passed compared to stations that the drivers had not passed yet during the sample period. Unsurprisingly, previously passed stations are more likely to be located closer to the driver's route. In particular, nearly all stations located directly on drivers' routes were passed previously by the driver. Conversely, stations that were not previously passed are more likely to be located further away from drivers' routes. Figure 3d shows the price distribution for passed and non-passed stations. Passed stations have prices that are on average \$0.03 higher compared to non-passed stations (also see Table A.4), and specifically stations in the left tail of the price distribution, with very low prices, are much less likely to be passed previously. This pattern suggests that the expenditure coefficients from a standard full information discrete choice model would be biased because some stations with lower prices may not actually be considered by drivers. In the next section we develop an empirical approach to estimate driver preference that is both tractable and accounts for imperfect information about stations and prices.

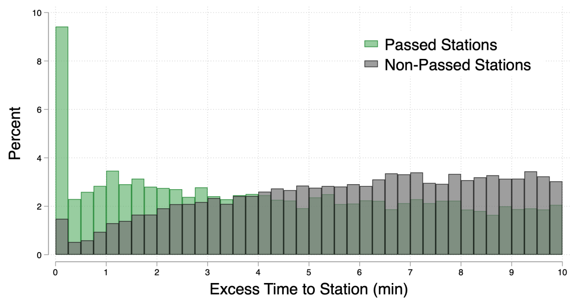
Figure 3: Characteristics of Chosen, Non-Chosen, and Previously Passed Stations



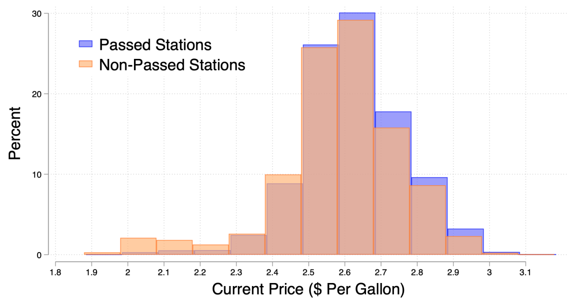
(a) Excess Time: Chosen vs. All Stations



(b) Excess Time: Chosen vs. Closest Station



(c) Excess Time: Passed vs. Non-Passed



(d) Price: Passed vs. Non-Passed

Notes: The histogram 3a shows the share of chosen stations that are at different excess times out of the way of the driver's quickest route in green and the share of all stations that are within 10 minutes of the driver's quickest route in gray. Histogram 3b shows the distribution of excess travel time for chosen stations in green and the distribution of excess travel time for the closest station to a driver's quickest route in tan. Histograms 3c and 3d show the distributions of excess time and prices for passed and non-passed respectively.

3 Model of Refueling Choice

Our model incorporates drivers' choice of whether to refuel, where to refuel, and the fuel purchase quantity. 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, and the outside option of not stopping for gas, which we denote $k = 0$. The utility that each driver receives for each station choice $k = 1, \dots, K$ is given by:

$$U_{ikt} = \alpha\, \mathbb{E}_i[p_{kt} \cdot q_{kt} | Z_{ikt}] + \gamma\, \text{ExcessTime}_{ikt} + X_k'\beta + W_{it}'\delta + \varepsilon_{ikt} \tag{1}$$

where $\mathbb{E}_i[p_{kt} \cdot q_{kt} | Z_{ikt}]$ is driver i's expected total expenditure (with unit price p_{kt} and quantity q_{kt}) associated with refueling at station k on the date of trip t. Throughout this section, the expectation operator $\mathbb{E}_i$ is understood to be taken over the driver's subjective distribution of stations' prices given observable information available at the start of the trip, Z_{ikt}. The driver's utility also depends on ExcessTime_{ikt}, the additional travel time to visit station k on trip t,[11] X_k, a vector of characteristics of the station k such as corporate brand, W_{it}, a vector of characteristics of the trip and driver including the amount of gasoline remaining in the fuel tank at the start of the trip and characteristics of the driver, and ε_{ikt} is an idiosyncratic preference shock. We normalize the utility the driver gets if they choose not to stop as $U_{i0t} = 0 + \varepsilon_{i0t}$.[12] Therefore, in this model, the driver simultaneously makes the choice of whether to stop on each trip and, if they stop, where to stop.[13] The unobserved error term ε_{ikt} has a generalized extreme value distribution for which ε_{ikt} may be correlated with ε_{ik^*t} ($k \neq k^*$ both greater than zero), while ε_{i0t} is uncorrelated with all other ε_{ikt}. This error structure matches that of the familiar nested logit model (Cardell, 1997) where all stations are in one nest, and we denote the nesting correlation parameter as $\lambda \in [0, 1]$.

In order to estimate Equation (1), we assume each driver makes refueling decisions as follows. At the start of each trip, the driver observes the information Z_{ikt} and the characteristics of the choice set, X_k and the trip W_{it}. The driver then uses these variables to form expectations over the expected fuel expenditure at each station, $\mathbb{E}_i[p_{kt} \cdot q_{kt}|Z_{ikt}]$. After forming these expectations, the driver observes their idiosyncratic preference shocks, ε_{ikt} and makes a choice over whether to stop and at which station to stop. The driver then proceeds to the station, observes the actual price at the station on that day and an additional idiosyncratic shock, η_{ikt}, that determines their true purchase quantity, and completes the purchase.

Unfortunately, we cannot observe each driver's expected fuel expenditures at each station at the start of each trip. To solve this problem, we start by partitioning the driver's information at the start of the trip into two components, $Z_{ikt} \equiv \{Z_{1ikt}, Z_{2ikt}\}$. The first component, Z_{1ikt}, includes variables that affect drivers' fuel purchase quantities, and the second component, Z_{2ikt}, contains variables that influence drivers' price perceptions. Using this notation, we then make two closely related assumptions about how drivers form expectations over purchase quantities and prices:

Assumption 1 *Drivers' expectations over the quantity of gasoline they would purchase at each station are a function of the vector Z_{1ikt}. Conditional on Z_{1ikt}, purchase quantities are independent of drivers' price expectations and the vector of unobserved idiosyncratic preference shocks, $\vec{\varepsilon}$.*

[11] Our calculation of the excess time out of the way to each station assumes that if the driver had made a different choice about where to stop for gasoline, they would still have traveled to that station from the same starting location and would travel to the same destination after leaving the station. This rules out a scenario, for example, where a driver chooses to pick up coffee at a different coffee shop depending on which gas station they stop at.

[12] This formulation of the utility of not stopping assumes that not stopping incurs a price expenditure of zero and a time driven out of the way of zero. Since we do not observe instances in the data where drivers run out of gas, we cannot recover the cost of running out of gas directly. Instead we allow this cost to be embedded in the increased value of stopping for fuel as the tank level decreases.

[13] Our static model rules out the possibility that consumers make dynamic decisions about where to stop for gas such as making a choice about whether and where to stop based on the characteristics of stations near the next trip.

Assumption 1 says that drivers' expectations over their purchase quantities will be dependent upon information such as current tank level, driver characteristics, and station characteristics, which is captured in Z_{1ikt}. This assumption is critical to our empirical approach because a driver who expects to purchase more gasoline—for instance because their tank level is low—can expect to save more money by driving out of their way to a cheaper station than a driver who expects to purchase only a small amount of gasoline. Assumption 1 further says that a driver's price expectation and idiosyncratic preference shocks are not allowed to directly enter into the driver's quantity expectation. Our assumption that gasoline purchase quantity is inelastic with regards to price is consistent with Houde (2012) and Hastings and Shapiro (2013), however, we do allow expected purchase quantity to depend on trip and station characteristics such as station brand and tank level. We also validate the assumption empirically in Appendix Table A.5 by regressing purchase quantities on prices, controlling for Z_1. This exclusion restriction aids identification of the full set of utility parameters, which we discuss further in Section 3.1.

Assumption 2 *Drivers' expectations over the price they would pay at each station are a function of the vector Z_{2ikt}. Conditional on Z_{2ikt}, drivers' price expectations are independent of drivers' purchase quantity expectations and the vector of unobserved idiosyncratic preference shocks, $\vec{\varepsilon}$.*

The second assumption means that drivers form expectations of the per-unit price they will pay at each station based on information like the long-run average price at that station, the price the last time they passed that station, and possibly the current price at the station. Our assumption is substantially weaker than the implicit assumption made in most discrete choice models that consumers are perfectly informed about current prices. Consequently, our modeling approach should reduce any bias that imperfect information would typically generate when estimating utility parameters using standard approaches. Assumption 2 explicitly rules out that these expectations of prices will be affected by characteristics of the driver's particular trip, like their current tank level. As discussed above, these characteristics are allowed to affect both the likelihood the driver stops on a given trip (via W_{it}) and the expected purchase quantity (via Z_{1ikt}).

We leverage these two assumptions to derive an empirically tractable formulation of the utility model allowing for imperfect information. More precisely, we express driver i's conditional expected fuel expenditure at station k on trip t as:

$$\mathbb{E}_i[p_{kt} \cdot q_{kt}|Z_{ikt}] = \mathbb{E}_i\left[p_{kt} \cdot \mathbb{E}_i\left[q_{kt}|p_{kt}\right]\middle|Z_{ikt}\right] = \mathbb{E}_i\left[p_{kt}|Z_{2ikt}\right] \cdot \mathbb{E}_i\left[q_{kt}|Z_{1ikt}\right], \tag{2}$$

where the first equality follows from the law of iterated expectations and the second equality applies Assumptions 1 and 2. The above expression makes clear that we can specify functional forms for drivers expectations of price and quantity separately.

We model each driver's expectation of the unit price at station k on trip t as the weighted sum of the station's current price p_{kt} and its long-run average price, $\overline{p}_k$, specifically $\mathbb{E}_i[p_{kt}|Z_{2ikt}] =$

$\theta p_{kt} + (1-\theta)\overline{p}_k$. The θ parameter is therefore the weight that information related to current prices plays in the formation of drivers' perceptions relative to information related to long-run average prices at that station.[14] Importantly, this specification nests the special case where $\theta = 1$ and drivers are perfectly informed about current prices, as well as the case where $\theta = 0$ and drivers only have information about long-run prices at stations. We also empirically explore whether replacing the long-run average price at the station with other variables, such as the price the last time the driver passed the station, provides a better fit of the model to the data.

We then additionally specify drivers' conditional expected purchase quantity as a linear function of individual, trip, and station characteristics:

$$\mathbb{E}_i[q_{kt}|Z_{1ikt}] = Z'_{1ikt}\phi. \tag{3}$$

Substituting these functional forms into the indirect utility function in Equation (1) we get:

$$U_{ikt} = \alpha \underbrace{\left(\theta p_{kt} + (1-\theta)\overline{p}_k\right)}_{\mathbb{E}_i\left[p_{kt}|Z_{2ikt}\right]} \cdot \underbrace{\left(Z'_{1ikt}\widehat{\phi}\right)}_{\mathbb{E}_i\left[q_{kt}|Z_{1ikt}\right]} + \gamma \text{ExcessTime}_{ikt} + X'_k\beta + W'_{it}\delta + \varepsilon_{ikt}. \tag{4}$$

We estimate Equation (4) in two steps. In the first step, we estimate $\widehat{\phi}$ by regressing fill quantities on the variables in Z_{1ikt}. The remaining parameters to estimate in the second step are the expenditure sensitivity α, time sensitivity γ, the weight on information related to current prices θ, preferences over non-price attributes β, the propensity of stopping to refuel δ, and the nesting parameter λ. We estimate these parameters by pseudo maximum likelihood and bootstrap the standard errors to account for the two-step estimation process.

Our two-step approach is similar in spirit to the common two-step approach for estimation of dynamic models (*e.g.* Hotz and Miller (1993)). Notably, drivers in our model use the price they expect to pay to form expectations of the total fuel expenditure if they stop at each station. However, when the driver actually arrives at the chosen station, they observe the true price on that day, p_{kt}, and the idiosyncratic error, η_t, and choose the fill quantity q_{kt} given that new information. Thus consistency of the two-step estimation approach hinges on the validity of Assumption 1. In particular, that drivers' form expectations about their fill quantity at each station prior to making a station choice. This assumption would be violated if drivers budget a fixed dollar amount for gasoline (*e.g.* \$20). In this case, the expected fill quantity would be mechanically related to the realized price at the station where the driver chooses to refuel. To provide support for our assumption, Figure A.6 in the appendix plots a histogram of the implied fuel expenditure across all refueling stops and shows that there is little evidence of

[14] We estimate $\overline{p_k}$ for each station using the mean price over the two-year sample we observe in the OPIS data.

a discrete jump in expenditure at any specific dollar amount.[15] Finally, our model imposes an implicit assumption that drivers are not engaging in a sequential search for gas stations. This assumption is supported by the fact that gas prices that drivers observed recently are uncorrelated with the decision to stop to refuel, conditional on tank level. Thus, the observed refuelling behavior is inconsistent with a sequential model of search (De los Santos et al., 2012).

3.1 Identification

To identify the expenditure coefficient α, we rely upon Assumptions 1 and 2, which imply that (expected) expenditure is not correlated with the idiosyncratic preference shock, ε_{ijt}. Here, Assumption 2 embodies the standard assumption that prices are uncorrelated with unobservable product quality. This is perhaps less of a concern for gasoline than for many other products since gasoline itself is a homogeneous product. However, price still could be correlated with the quality of the gas station's non-gasoline characteristics such as its location or convenience store. Our ability to observe each driver's excess time to reach each station on their current trip gives us the unique ability to control for the quality of a station's location as viewed by each driver. To account for the correlation between price and station attributes such as the quality of the convenience store, we include gas station brand fixed effects in the utility function.[16]

In our context, however, we need to also assume (via Assumption 1) that the expected quantity of gas purchased at each station is uncorrelated with the idiosyncratic preference shock, once we condition on Z_{1ikt}. To ensure this assumption holds, we include every variable that affects purchase quantity via Z_{1ikt} directly in the utility function (e.g. in X_{kt} or W_{it}). To see why this is necessary, suppose we allow gas brand to affect purchase quantity via Z_{1ikt}. Then, if a driver expects to purchase more gas at Costco, this will increase their expected expenditure at Costco as well as their utility, since they will acquire more gas during such a stop. Thus, the Costco fixed effect in the utility function captures this extra gas purchase, preventing it from entering the idiosyncratic error term and biasing our estimate of α.

Assumption 1 further implies that the expected purchase quantity can be identified in an initial regression of observed purchase quantities on the driver's information, Z_{1ikt}. The assumption rules out that there is anything beyond Z_{1ikt} that affects the expected purchase quantity and is correlated with preference shocks.[17]

[15] There is a small increase in the likelihood of expenditure around \$20, but this increase would imply that only very small percentage of refueling stops are affected by this type of budgeting.

[16] In the following section, we show evidence that this assumption is likely to hold in our setting, and that controlling for additional station characteristics doesn't change our results.

[17] This rules out, for instance, a situation where drivers know that they will purchase more than their standard amount of gasoline if they stop at a station with a particularly high ε_{ikt}. Our preferred specification includes gas brand dummies in Z_{1ikt}, so a violation of this assumption would only occur if drivers expect to purchase more gasoline at a specific station relative to other stations with the same brand.

The weight that drivers place on information related to current price, θ, is identified through within-station price variation over time. As an example, consider a Costco station that sets a relatively low long-run average price relative to other competing stations. On days that the Costco station sets price at its long-run average price ($\overline{p_k} = p_{kt}$), any $\theta \in [0, 1]$ implies the same choice probability. Thus, the station's choice probability on a day that $\overline{p_k} = p_{kt}$ allows us to identify α and the other utility parameters just as in a standard discrete choice setting. Now consider a day when the Costco station sets price higher than its long run average price but all other competitors keep prices the same; θ is identified by the extent that drivers substitute away from choosing the Costco station on that day. If drivers do not change their choices at all, then θ must equal one, but if Costco loses market share that implies $\theta < 1$. Assumption 1 guarantees that a change in Z_{2ikt}, Costco's current price in this example, does not affect the driver's expected quantity of gas, conditional on Z_{1ikt}. This exclusion restriction therefore provides important variation for identifying θ. Finally, the identification of the remaining utility parameters β, δ, and λ follow standard arguments.

4 Results

We first present the results of our refueling choice model. We then use our estimates to calculate drivers' implied value of time and their potential gains from improved information about gas stations and current prices. Finally, we explore heterogeneity in driver preferences and the robustness of our results to changes to the model specification and underlying assumptions.

4.1 Model Estimates

We report the estimates from the first step "fill quantity regressions" in Table 2. In these regressions, we allow the expected fill quantity to vary as quadratic function of the drivers' tank level at the start of trip. In our preferred specification, Column (4), we also allow the expected fill quantity to vary across gas station brands. The regressions show that the expected fill quantity increases by approximately a half a gallon for each one unit reduction in the driver's initial tank level. The intercept of regression in Column (4) implies that drivers' expected purchase quantity at an unbranded station (or small brand) is approximately 9 gallons if their tank is empty at the start of the trip. We find that most gas brands are not associated with a statistically significant change in expected purchase quantity—notable exceptions are Costco, Citgo and Speedway which are associated with higher expected purchase quantities relative to unbranded stations, and Marathon which is associated with lower purchase quantities.

Table 3 presents estimates for several specifications of our indirect utility model. Namely, the first three columns of Table 3 show the results for specifications where drivers are assumed to be fully informed about current prices at all stations in their choice set. Columns (4) through (6) of Table 3 show results for specifications that allow for imperfect information, in

Table 2: Fill Quantity Regressions

	Fill Quantity (Gallons)			
	(1)	(2)	(3)	(4)
Intercept	9.121	9.172	8.882	8.854
	(0.3201)	(0.3977)	(0.3437)	(0.4268)
Tank Level	-0.5138	-0.5214	-0.4159	-0.4226
	(0.1902)	(0.1883)	(0.2173)	(0.2148)
(Tank Level)^2	0.0072	0.0083	-0.0130	-0.0125
	(0.0230)	(0.0228)	(0.0278)	(0.0275)
Observations	865	865	744	744
R^2	0.04387	0.09119	0.05085	0.10267
Station Brand Fixed Effects	N	Y	N	Y
Choice Set	All	All	Passed	Passed
Models	1,4	2	5	3,6

Notes: The dependent variable is the imputed fill quantity associated with each observed refueling stop. In the models with brand fixed effects, the intercept represents the fill quantity at zero tank level for small brands and unbranded stations. The regressions estimates are used to predict expected fill quantity conditional on stopping for each trip on covariates. The predictions are used as an inputs for the choice models presented in Table 3. The first regression is used to predict expected fill quantity for Models (1) and (4) in Table 3, the second regression is used for Table 3 Model (2), the third regression for Table 3 Model (5), and the fourth regression applies to Table 3 Models (3) and (6).

which drivers' price perceptions are allowed to vary with either the current price or the long-run average price (or both). The first, fourth, and fifth columns show specifications without station brand fixed effects and in the remaining Columns (2, 3, and 6) we add station brand dummies for the ten most commonly chosen gas brands, with the remaining stations grouped into a separate generic brand category. Finally, in Columns (3, 5, and 6), we restrict the choice set to include only stations that the driver has previously passed, and therefore drop choice situations where the chosen station had not previously been passed. Our preferred specification is Column (6) which allows for imperfect price information, includes gas brand fixed effects, and restricts the choice set to previously passed stations.

The top section of Table 3 demonstrates how the value of stopping to refuel depends on the amount of fuel remaining in the tank. We see that stopping becomes more valuable as the driver's tank level falls. Moreover, the quadratic term on tank level is negative in all specifications implying that the probability of stopping is increasing at an increasing rate as the tank level declines. Figure A.5 in the Appendix plots the empirical probability of stopping across tank levels and shows that when tank level is over 50% the probability of stopping is close to zero, but the probability rises non-linearly as the tank falls below 25%.

The second section of Table 3 reports the parameters that determine the station choice conditional on stopping. For the specifications that allow for imperfect price information,

Table 3: Driver Preference Estimates

	Full Price Information			Imperfect Price Information		
	(1)	(2)	(3)	(4)	(5)	(6)
Decision to Stop						
1[Stop] $\times$ Constant	1.690	0.753	0.767	7.745	7.075	6.414
	(0.357)	(0.374)	(0.358)	(0.744)	(0.697)	(0.845)
1[Stop] $\times$ Tank Level	-0.298	-0.244	-0.182	-0.644	-0.483	-0.453
	(0.042)	(0.044)	(0.041)	(0.055)	(0.048)	(0.052)
1[Stop] $\times$ (Tank Level)^2	-0.037	-0.038	-0.052	-0.031	-0.060	-0.060
	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)
Station Choice						
α - Expenditure	-0.172	-0.136	-0.138	-0.429	-0.409	-0.384
	(0.015)	(0.015)	(0.015)	(0.033)	(0.032)	(0.038)
θ - Weight on Current Price				0.168	0.172	0.169
				(0.045)	(0.047)	(0.048)
γ - Excess Time (minutes)	-0.202	-0.229	-0.168	-0.216	-0.160	-0.176
	(0.013)	(0.018)	(0.016)	(0.015)	(0.012)	(0.016)
Nesting Parameter						
λ	0.358	0.406	0.337	0.378	0.317	0.350
	(0.025)	(0.031)	(0.031)	(0.025)	(0.023)	(0.031)
Station Brand Fixed Effects	N	Y	Y	N	N	Y
Number of Stops	865	865	744	865	865	744
Number of Trips	22114	22114	21355	22114	21355	21355
Observations	2663807	2663807	815070	2663807	815070	815070
Choice Set	All	All	Passed	All	Passed	Passed
Implied Value of Time ($/hour)						
	70.51	101.24	73.37	30.19	23.42	27.54
	(7.06)	(14.00)	(9.74)	(2.18)	(1.67)	(2.45)

Notes: Table reports pseudo maximum likelihood estimates of driver preferences. The expected fill quantities are predicted from the regressions in Table 2. The full information models assume drivers know current gas prices at each station and the imperfect information models allow drivers' price perception to be a weighted average of current price and station average price. λ is the nested logit correlation parameter. In models with brand fixed effects, the constant in the decision to stop represents small brand and unbranded stations. Choice Set = "All" indicates that all stations with 20 minutes of the driver's route are included in the choice set. Choice Set = "Passed" means stations that the driver has previously passed that are within 20 minutes of the route are included in the choice set. Bootstrapped standard errors are reported in parentheses. The implied value of time (per hour) is calculated as $60 \cdot \frac{\gamma}{\alpha}$.

we estimate that drivers place a relatively low weight on current prices (16%-17%) relative to long-run average prices when forming price perceptions. In other words, drivers respond more than four times more to long-run variation in prices across stations compared to day-to-day variation in station-level prices. 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.

The gas brand coefficients (not reported) show that Costco and Meijer are the brands most likely to be chosen after controlling for price and location. Smaller brands (outside the ten largest brands) are the least likely to be chosen conditional on price and location. Interestingly, we find that the estimated expenditure coefficients are slightly smaller in magnitude after we control for gas brand fixed effects. Typically, we would expect brand controls to increase the magnitude of the expenditure coefficient if brand quality is positively correlated with price. However in the gasoline market, high quality firms may set lower gas prices as a "loss leader" strategy to attract customers to visit their convenience stores or grocery stores.[18] The estimates presented in Table 3 suggest that the bias from price endogeneity is less problematic compared to the bias caused by mispecification of consumer information in this setting. However, to the extent that price is still correlated with unobservables, we expect our estimates would be more likely to overstate the magnitude of the expenditure coefficient due to the apparent negative correlation between price and station quality. Thus, our model should provide a conservative estimate of drivers' value of time which we discuss in the next subsection.

4.2 Value of Time

By taking the ratio of the estimated coefficients on excess time and expected expenditure, we obtain estimates of drivers' *value of time* (VOT). The VOT provides a measure of driver preferences that is both easy to interpret and important for policymaking. Intuitively, the VOT is determined by the marginal rate at which drivers trade off time savings—by selecting more conveniently located stations—and expected dollar savings at the pump. Specifically, we calculate the VOT as follows:

$$\text{VOT} = 60 \times \frac{d\,\mathbb{E}_i[p_{kt} \cdot q_{kt}]}{d\text{ExcessTime}_{kt}} = 60 \times \frac{\partial U_{ijt}/\partial \text{ExcessTime}_{kt}}{\partial U_{ijt}/\partial\,\mathbb{E}_i[p_{kt} \cdot q_{kt}]} = 60 \times \frac{\gamma}{\alpha}. \tag{5}$$

Here, we multiply by 60 to convert the value of time from dollars per minute to dollars per hour. The first two specifications, in which we assume drivers consider all stations and know

[18] 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 could also explain why these brands are preferred after controlling for (list) price and location. We do not observe which customers have discount cards in our data.

all stations' current prices, imply distinctly high values of time. We estimate a VOT of $71 per hour in the specification without gas brand fixed effects and $101 per hour after including brand fixed effects.

These VOT estimates will be biased upwards (the expenditure coefficient will be biased towards zero) if consumers are not aware of all stations and their respective prices. To account for the possibility of limited consideration sets (Abaluck and Adams-Prassl, 2021; Goeree, 2008), we take advantage of our detailed information on drivers' geospatial locations to identify stations that each driver passed within our sample period and therefore are more likely to informed about. The specification in Column (3) assumes that drivers only consider (and are informed about) the set of stations that they have *previously passed* and are within 20 minutes of their route.[19] This specification, which also includes brand fixed effects, yields a 28% lower value of time of $73 per hour relative to Column (2).

Our VOT estimates fall substantially further after we allow for drivers to be imperfectly informed about current gas station prices. In specification (4), we include all nearby stations in the driver's consideration set but relax the assumption that drivers must know all current prices, which implies a VOT estimate of $30.19 per hour. The VOT estimate drops further to $23.42 per hour in Column (5) when we add gas brand fixed effects. Comparing these results to Houde (2012)—which finds a very high value of time of $54 per hour in a model with perfect information—and our earlier estimates suggests that properly accounting for drivers' information set is important for recovering unbiased estimates of drivers' VOT.

Our preferred VOT estimate of $27.54 per hour is in Column (6). In our preferred model we allow for imperfect price information, include brand fixed effects, and limit the consideration set to only previously passed stations. Comparing the VOT in the first three columns to the last three columns of Table 3 makes clear the importance of modelling imperfect information. In particular, our θ estimates imply that drivers respond very little to deviations in current prices across stations relative to differences in long-run price differences across stations. Intuitively, if drivers are not aware of changes in current prices, we will see relatively little substitution towards stations that reduce prices on any specific day. Thus, we would understate drivers' willingness to drive further to save on gas expenditures if we assume that they know all of these (current) prices. Therefore, our specifications that allow drivers' price perceptions to deviate from current prices leads to lower and more credible estimates of the VOT

Our preferred estimate of drivers' VOT, $27.54 per hour, is somewhat higher than the recent estimates by Goldszmidt et al. (2020) of $19.38 per hour. Noteably though, our estimates are based on drivers' value of driving time, whereas Goldszmidt et al. (2020) measure the

[19]Recall from Table (1) that drivers stop at stations that they have previously passed nearly 90% of the time. The specification in Column (3) discards the stops at previously un-passed stations from the analysis. Figure (3) illustrates why including the full set of stations in drivers' choice sets would likely cause bias; previously passed stations are more likely to be closer to the driver's optimal route and also have higher average prices compared to non-passed stations. Therefore, drivers may fail to choose cheaper stations that are further from their route because they are not aware of these stations, which would in turn lead us to overestimate drivers' value of time.

value of waiting time savings on a ride hailing platform.[20] Our estimates are also higher than the VOT currently used by U.S. government agencies, which ranges from 33% to 50% of the wage rate.[21] For our sample, US government guidelines would imply a VOT between $10.21 and $15.47 per hour. In contrast, our VOT estimate of $27.54 amounts to 89% of the median household income for the census tracts in which our drivers live.[22] Thus, our estimates are suggestive that US lawmakers may be undervaluing time-saving investments and regulations.

A relatively high estimate of the VOT has a numerous policy implications. For example, investment in infrastructure that reduces travel time—such as highway expansions and improvements—would yield larger benefits to drivers than would be implied using lower VOT estimates such as those currently used by policymakers. Additionally, a high VOT indicates that fuel economy standards over the past decade may create added benefits to drivers to the extent that these policies reduce the amount of refueling stops that drivers have to make. As a final example, drivers' value of electric vehicles will be highly dependent upon the number of charging stations available and the charging speed of those stations. We investigate this third example in more detail in Section 5.

4.3 Value of Information

We next use our model to evaluate the consumer welfare effects of imperfect information in the refueling market. Our previous estimates show that drivers are not perfectly informed about current prices and respond much more strongly to stations' long-run average prices. In addition, the empirical evidence suggests that drivers are not aware of all stations that are available nearby their route when making refueling choices. Therefore, we calculate the gains to drivers from obtaining better information about available gas stations and prices, which is relevant for assessing the market impacts of government price disclosure requirements or price reporting websites and apps (Byrne and de Roos, 2017; Byrne et al., 2015; Lewis and Marvel, 2011; Luco, 2019).

We measure the welfare effects of changing drivers' information using a similar framework to Leggett (2002), Allcott (2013), Schmeiser (2014), Train (2015) and Houde (2018). Under this framework, consumers make purchase decision based on imperfect *perceptions* about product attributes, and then after making the purchase choice, ex-post utility depends on *actual* product attributes. In our setting, drivers are imperfectly informed about stations' prices when they choose a refueling station, but then they must pay the current posted price when they arrive at their chosen station. Thus, better information about current prices can increase

[20]It is plausible that individuals value time savings differently across transportation modes, for example, individuals may able to easily complete other tasks while waiting or while traveling in an ride share vehicle relative to driving their own vehicle.

[21]For example, the Environmental Protection Agency uses a VOT of 33% of the wage (Cesario, 1976), and the Department of Transportation use a VOT of 50% of the wage rate (Small et al., 2005; White, 2016).

[22]To calculate median wages, we first take the average of median census-tract incomes where the drivers live. Then we follow the the U.S. DOT (White, 2016) and divide annual income by 2010 hours worked per year.

ex-post utility through changing station choices, and therefore, the actual price paid for fuel.

To understand the calculation of consumer surplus when drivers have imperfect price perceptions, consider a driver facing two choice scenarios $s \in \{0, 1\}$. The two choice scenarios can differ due to the set of stations in the driver's choice set and because the driver may have different perceptions about stations' prices across the scenarios. In both scenarios, drivers select a station based on *perceived* expenditure but ex-post utility is determined by the *actual* expenditure. Let P_j be the actual expenditure to refuel at station j, but if $s = 0$, the driver perceives the expenditure to refuel at station j as P_j^{0*} and if $s = 1$, the driver perceives the expenditure as P_j^{1*}.[23] Accordingly, the driver's expected utility from choosing j and therefore the probability of choosing j depends on the driver's perception of expenditure. We denote the perceived utility from choosing j in scenarios s as v_j^{s*} and the probability of choosing station j as π_j^{s*}.

We calculate the change in expected consumer surplus for driver i on trip t between scenario 1 and scenario 0 as follows:

$$\Delta CS = -\frac{1}{\alpha}\left[\ln\Big(1 + \big(\sum_{j\in C^1} \exp(\frac{v_j^{1*}}{\lambda})\big)^{\lambda}\Big) - \ln\Big(1 + \big(\sum_{k\in C^0} \exp(\frac{v_k^{0*}}{\lambda})\big)^{\lambda}\Big)\right]$$
$$- \Big(\sum_{j\in C^1} \pi_j^{1*}(P_j - P_j^{1*}) - \sum_{k\in C^0} \pi_k^{0*}(P_k - P_k^{0*})\Big). \qquad (6)$$

Where C^0 and C^1 are the choice sets for scenarios $s = 0$ and $s = 1$, respectively. The first line of Equation (6) is analogous to the standard formula for a change in consumer surplus for the nested logit model (Small and Rosen, 1981), except that in this setup, these terms calculate the change in drivers' ex-ante *perceived* consumer surplus. The terms on the second line adjust consumer surplus to account for the possibility that the driver may end up paying prices that differ from their initial perceptions.

We apply Equation (6) to find the gain in consumer surplus from fully informing drivers about current prices for each trip in our data. In particular, the first row of Table 4 shows the change in consumer surplus if drivers became fully informed about current prices (i.e. P_j^{1*}=P_j) relative to the baseline case where drivers perceive prices as a weighted sum of current price and long-run average price, with the weights determined by $\widehat{\theta}$. For this first calculation, we hold the choice set fixed as the set of previously passed stations. We find that drivers learning current prices for previously passed stations would only slightly improve consumer surplus by 8.4¢ per refueling stop (1.1¢ per gallon purchased). In the second row of Table 4, we calculate the change in consumer surplus from adding all gas stations within 20 minutes of drivers' routes to their choice sets, relative to the baseline case where drivers only consider stopping at stations that they have previously passed. For this calculation, we hold constant drivers'

[23] We assume that the actual prices at each station are held fixed across the two choice scenarios.

imperfect perceptions of prices . We find that adding these stations to the choice set similarly increases consumer welfare by 5.8¢ per stop (0.8¢ per gallon purchased). Finally, in the last of row of Table 4 we estimate the gains from both informing drivers about all nearby stations and about current prices at all these stations. In this case, the consumer welfare benefits remain small in magnitude—14.5¢ per gas stop or 1.9¢ per gallon purchased.

Table 4: Value of Information

	Δ Consumer Surplus vs. Baseline ($)	
	Per Gallon	Per Stop
Current Prices Known, Choice Set = Passed Stations	0.011	0.084
Imperfect Price Info, Choice Set = All Stations	0.008	0.058
Current Prices Known, Choice Set = All Stations	0.019	0.145

Notes: Each cell shows the normalized change in consumer surplus (CS) from a change in drivers' information about current prices and/or nearby stations. We calculate the change in CS relative to the baseline case where drivers only consider previously passed stations and are imperfectly informed about current prices. In the baseline case, station price expectations are a function of $\hat{\theta}$. We first calculate the expected change in CS for each trip in the data. We then sum the expected change in CS across all trips. Finally, in each column, we divide the aggregate change in CS by the total gallons purchased, and total stops, respectively.

Why do consumers benefit so little from better information? Our model estimates indicate that drivers place a high value on their time, which means they strongly prefer to avoid traveling far from their routes to refuel. Figure 2 shows that on an average trip, drivers could save roughly \$0.64 (\$0.08/gal. $\times$ 8 gallons purchased) by finding the cheapest station within two minutes of their route relative to stopping at a random station directly along their route. However, if the cheap station is located two minutes away from the route, the time cost associated with visiting the cheapest station would be \$0.92 (\$27.54/hour $\times \frac{2}{60}$ hours). Thus, most drivers are unlikely to make substantially different station choices when they learn more about new stations or about current prices. Figure A.7 in the Appendix also provides additional intuition about the important channels that contribute to our value of information estimates. More specifically, we plot the change in consumer surplus from adding incremental stations to drivers' choice sets relative to the baseline. Drivers' baseline choice sets include an average of 37 previously passed stations that lie within 20 minutes of their optimal route. In the figure, we plot how consumer surplus per gas stop changes as we sequentially add unpassed stations into the choice sets, starting with stations that are nearest to drivers' routes in terms of excess time. The figure confirms that drivers benefit the most from learning current prices at stations very close to their route, which allows them to reduce fuel expenditures without significantly increasing travel time. Whereas, the marginal value of considering additional stations sharply decreases for stations further away from the route.

Noteably, our value of information results rest on the assumption that stations do not update their pricing decisions in the counterfactuals. Thus, our calculation is perhaps more appropriate for assessing the benefits of offering a single driver or a small set of drivers additional information about prices or nearby stations. Luco (2019) shows that price disclosure can increase prices by facilitating collusion, which suggests that our estimates could yield an upper-bound on the consumer welfare benefits of large-scale price disclosure policies.

In sum, our estimates suggest that consumers stand to benefit relatively little from government pricing disclosure requirements or from informational websites like *Gas Buddy*. In contrast, price disclosure can potentially harm consumers if that information is used to facilitate firm collusion (Byrne and De Roos, 2019; Luco, 2019).

4.4 Heterogeneity, Sensitivity, and Robustness Analysis

The disaggregate nature of our data allows us to further explore the heterogeneity in preferences across different types of drivers in our sample and across different types of trips. Appendix Table A.6 shows the results for specifications that allow both the expenditure coefficient, α, and the weight on current prices, θ, to vary by age, gender, and census tract income. Additionally, we allow for heterogeneity in price sensitivity across the time of trip (weekday versus weekend).[24]

We find evidence of substantial heterogeneity in expenditure sensitivity across age groups. Column (1) indicates that the oldest drivers (age 60-70)—who are presumably more likely to be retired—are substantially more price sensitive relative to younger drivers. According to Column (2), women appear to be more price sensitive than men. Column (3) allows price sensitivity to vary across driver income groups and Column (4) allows for heterogeneity across weekend versus weekday trips, but we do not find statistically significant differences in price sensitivity across income or trip time. The demographic interactions with θ are noisy but we find some evidence that women and older drivers are relatively less responsive to current prices. In particular, we estimate that the weight that both women and drivers aged 60-70 place on current prices is not statistically distinguishable from zero.

For ease of interpretation, we also report the average marginal effects corresponding to each of the demographic group indicators on the value of time in Appendix Table A.7. We report the marginal effects for the model which includes all the demographic interactions within a single specification (Column (5)). The interactions of the demographics with expenditure have less statistical power when we include all of them together in the same model but the estimates remain suggestive of sizeable heterogeneity in the value of time across demographics. The point estimates show that a driver being in the oldest age category (age 60-70) is associated with a \$11/hour (39%) reduction in the VOT. Similarly, women's implied value of time is \$5/hour (18%) lower than men's. In addition, we see that drivers from high-income census tracts have a \$17/hour (72%) higher VOT compared to drivers in middle-income census tracts. In sum, the heterogeneity estimates—although statistically imprecise—are generally in line with demographic VOT patterns that we would expect.

[24]For the "fill quantity regressions" we fit a flexible function that interacts both tank level and tank level squared with dummy variables for each of the demographic groups or trip types (*e.g.* weekend).

Having established our baseline and heterogeneity analyses, we also perform a series of robustness checks to support the validity of our VOT estimates. There are several potential issues that could lead to biased estimates of the model parameters: (1) misspecification of drivers' consideration sets, (2) misspecification of drivers' price expectations, and (3) correlation between prices and the error term in utility. Accordingly, we estimate several sets of alternative models to explore whether and how each of these issues might bias our estimates.

Our first set of sensitivity checks examines how changing our assumption about drivers' choice sets affects our estimated value of time. Recall, that our specification of price perceptions requires that drivers be informed about either current price or long-run average price for each station in their choice set. To increase the plausibility of this assumption, we only include stations that drivers have previously passed in the choice set under our preferred specification. However, it is plausible that drivers may be unaware of both long-run average prices and current prices at stations they have passed previously.

Appendix Table A.8 explores how altering the sample or reformulating the choice set changes the parameter estimates and the VOT. In the second and third column, we continue to assume the choice set includes only previously passed stations within 20 minutes of the route, but we further restrict the estimation sample to include only trips that originated nearby the drivers' homes. The motivation for these specifications is that drivers may be better informed about station prices in areas where they drive frequently. We find that restricting the sample to trips within ten or five miles of home imply a value of time of \$23.98/hour and \$28.40/hour—both which are very similar to—and not statistically different from—our baseline estimate of \$27.54/hour. In the fourth and fifth column, we restrict the choice sets to include only stations that the driver has passed within the past 14 days and past seven days respectively. Again, we obtain very similar estimates of the VOT—\$26.47 and \$23.53, respectively. These robustness checks provide evidence that altering our assumption about driver's station information (*i.e.* choice set) does not substantially impact our value of time estimates.

Next, in Appendix Table A.9 we test the sensitivity of our estimates to alternate formulations of drivers' price perceptions. Our estimates of the expenditure coefficient will also be biased if we mis-specify drivers' perceptions about station prices. In our base model, we assume that drivers' price perceptions can be expressed as a function of a station's current price and a station's long-run average price. However, it is reasonable to believe that drivers may use different information to form price perceptions. Notably, drivers could incorporate *recent* prices posted by a station in forming price expectations. The second column in Table A.9, we allow the price perception to be a function of the price posted the last time the driver passed a station, as well as the current price. In Columns (3, 4, and 5) we allow price perceptions to be a function of current price p and $\overline{p}$, where $\overline{p}$ is defined as the mean price at the station over the current month, current quarter, and current half-year, respectively. We find that estimates are somewhat sensitive to the selected variable that enters expectations. More specifically, we find that the specifications that include only more recent price information, such as price most recently passed and mean price over the current month, imply higher values of time between $71 and $60 per hour. The specifications where perceptions are function of the station's quarterly mean price or semi-annual mean price imply VOTs of $38-41 per hour, closer to our baseline estimate. Despite the sensitivity of the value of time estimates across the candidate variables entering the price perception function, a clear pattern emerges. Namely, we see that allowing for longer-run price information in the expectation function implies lower values of time. Thus, our baseline model provides a relatively conservative estimate of the value of time. Moreover, by comparing the likelihood function across these alternative models, we see that our baseline model, which includes current and long-run average price (over the full sample), provides the best fit to the data.[25]

Finally, Appendix Table A.10 illustrates the robustness of our estimates to adding more station-level control variables in the utility function. An important factor that may influence station choices and also be correlated with station prices is the stations' locations. Our baseline specification includes information about the relative distance to reach each station from the driver's route. However, drivers may also prefer to stop in certain types of neighborhoods even after conditioning on excess time to reach each station. Accordingly, we collect data on stations' neighborhood characteristics at the census-tract level from the 2010 American Community Survey. More specifically, we add controls for census tract median income and the population density where each station is located. The point estimates in Table A.10 suggest that drivers may slightly prefer stations located in higher income and lower population density areas, although the coefficient estimate census-tract income is not statistically significant. Moreover, we see that our value of time estimates do not change much after adding these neighborhood controls ($27.28-$27.30 per hour) and are not statistically distinguishable from our baseline estimate.

5 Implications for Transportation Electrification

Currently, the vast majority of on-road vehicles are powered by burning fossil fuels such as gasoline and diesel. Transitioning away from gasoline and diesel towards EVs offers a promising way to reduce greenhouse gas emissions, so long as the electricity used to charge these vehicles comes from low-carbon sources. However, such a transition would entail large changes to when, where, and how drivers refuel. Thus, the welfare impacts of transportation electrification will depend in part on the value of time and also whether drivers lose or save time as result of switching to an EV. Our estimates of driver preference and the VOT can help elucidate the relationship between refueling infrastructure and EV adoption. Namely, our empirical results show drivers' VOT is roughly equal to the average wage rate, which is

[25] We also tried specifications that allowed the price perception function to depend on more than two variables, but we found that these specifications either had extremely large standard errors or did not converge due to the high degree of correlation of station-level prices over time.

markedly higher than the standard VOT used in U.S. policy-making of 33-50% of the wage rate. In this section, we discuss the implications of our refueling preference and VOT estimates in the context of transportation electrification.

5.1 Drivers with Home Charging

For many households, purchasing an EV can provide added convenience by allowing drivers to refuel at home instead of traveling to fueling stations. In particular, drivers with access to a garage or carport with electric charging can plug in their EV when they arrive at home and charge the vehicle overnight. By allowing drivers to charge at home, EVs effectively reduce the time cost that drivers spend visiting fueling stations in a gasoline-powered car.

We use our VOT estimates along with our data on driver refueling behavior to calculate the gains from avoiding travel to gasoline stations over the lifetime of an EV. In particular, we calculate the total time savings from home charging relative to gasoline refueling as follows:

$$\text{Value of Time Saved} = \sum_{t=0}^{25} \frac{1}{(1+r)^t} \times M_t \times \frac{\text{Gas Stops}}{\text{Mile}} \times \frac{\text{Excess Time}}{\text{Gas Stop}} \times \text{VOT} \tag{7}$$

where r is the annual discount rate, which we assume to be 0.05, and M_t is the survival-weighted mileage driven in year t of the vehicle's life given by Lu (2006). We calculate this value of time saved separately for cars and light trucks following Lu (2006) and then assume a 70% market share of light trucks following the April 2019 NADA "Market Beat."[26] We calculate both the number of gasoline stops per mile drive from our data. We calculate the refueling time per gasoline stop as the average excess time per stop in our data (1.78 minutes) plus the waiting time required to pump the purchased gasoline assuming a gas pump rate of 10 gallons per minute (0.76 minutes).[27] Applying our preferred value of time estimate of \$27.54 per hour implies that drivers would value refueling at home at \$829 over the lifespan of an EV. The majority of the value, \$581, comes from avoiding driving time to gas stations and the remaining \$248 derives from avoiding waiting time at the gas pump.

This calculation highlights an important benefit of switching to an EV for drivers that have access to home charging. However there are several important caveats to take into account when interpreting the results. First, our calculation abstracts away any benefits that consumers may obtain from visiting gas station convenience stores. Second, our calculation does not account for potential welfare consequences of having to recharge an EV on longer road trips. Finally, the baseline time savings estimate masks substantial heterogeneity in the

[26] https://blog.nada.org/2021/05/05/nada-market-beat-new-light-vehicle-sales-top-18-million-unit-saar-for-second-straight-month/

[27] The EPA regulation require that gasoline pumps cannot operate at a rate above 10 gallons per minute. We chose not use the observed time that the drivers spend at the gas station to avoid restroom breaks and convenience store visits as refueling time.

potential time savings across drivers. Importantly, aggregate time savings will depend on how much each individual drives and refuels their vehicle. For example, we observe some drivers who stop for gasoline upwards of five times per week which would imply aggregate time-savings of over $2,500 over the lifespan of an EV. Similarly, our estimates of the heterogeneity in drivers' value of time suggest that working-aged drivers or high-income drivers may obtain more value from charging at home than others.

5.2 Drivers without Home Charging

While drivers with the ability to charge at home could generally save on refueling time by adopting EVs, there are many drivers who do not have a dedicated space for home charging (Traut et al., 2013). Drivers without home charging capability will need to travel to shared charging station infrastructure to refuel. Therefore, the time cost associated with refueling an EV for these drivers will depend on both the density of charging network (e.g. the number of charging stations) and the speed of the charging technology. Concerns over the ability of drivers to charge away from home have lead to substantial policy interest in expanding public charging infrastructure. For example, President Biden's *Infrastructure, Investment, and Jobs Act* includes funding for 500,000 charging stations nationwide.

In this section, we use our estimates of drivers preferences to better understand four important questions relevant to public EV charging infrastructure investment. First, given the current EV charging network, how much more time would drivers spend refueling an EV relative to a gasoline vehicle? Second, how has this refueling time differential changed with the expansion of the EV charging network from 2011-2021? And third, how large are the time-saving benefits to drivers from investments in the public charging network? In particular, we assess the marginal value of increasing the speed of charging stations (e.g. investing in more DC fast chargers) relative to the marginal value of increasing the number of available charging stations (which likely charge at lower speeds).

To answer these questions, we combine our geospatial data on drivers' trips with data from the U.S. Department of Energy (DOE) Alternative Fuels Data Center that reports the locations and entry dates of U.S. EV charging stations. Public charging has expanded substantially in the last decade, Table 5 shows that the number of charging stations in Michigan and Ohio expanded over 30-fold from 46 stations in 2011 to 1,601 stations in 2021. Moreover, the share of DC faster chargers in the network increased from 1% to 17%. But despite this expansion, there remains far fewer EV charging stations compared to gas stations. Appendix Table A.11 compares the number of gas stations to the number of electric charging stations that are located near drivers' routes in our data. On an average trip in our data there were 37.17 gas stations (23.41 previously passed stations) within five minutes of drivers' optimal routes but only 4.75 EV charging stations. Moreover, on a typical trip, a driver could stop for gas within a one minute deviation of their route but the closest EV charger was located over four minutes away.[28]

Table 5: Excess Refuel Time as Function of Station Density and Charge Speed

	# of Stations	Fast Charger Share [0, 1]	Mean Charger Speed (kW)	Excess Time (minutes)	Excess Time Elasticity w.r.t. # of Stations	Excess Time Elasticity w.r.t. Charger Speed
2011	46	0.01	17.81	121.64	-0.38	-0.69
2016	326	0.06	21.98	51.42	-0.23	-0.61
2021	1601	0.17	29.02	30.63	-0.11	-0.52

Notes: The first column shows the number of total EV charging stations located in Michigan or Ohio for each year as determined by each station's entry year in the DOE data. The second column shows the share of stations that offer DC fast charging by year, if a station offers both DC fast chargers and AC chargers, we code that station based on its share of fast chargers (e.g. 0.5 if half the chargers at the station are fast chargers). The third column calculates the mean estimated charging speed across all stations in the network. The fourth column shows our estimate of excess time per refueling stop based on that year's network configuration. The last two columns show the the elasticity of excess time with respect to changing the number of stations and charger speed, respectively.

Beyond understanding the distribution of charging station locations relative to drivers routes, we use our refueling choice model to predict when and where drivers would choose to refuel an EV and to measure the time costs of refueling an EV compared a gasoline vehicle. In particular, we reconstruct each driver's refueling choice set for each trip with the locations of EV charging stations within a 20-minute (driving) deviation from the driver's optimal route. When making their EV refueling decision, we allow drivers to either (1) drive to the charging station and wait for their vehicle to recharge or (2) park their vehicle at the charging station and walk to their destination, assuming that drivers' walking speed is 3 miles per hour. For example, a driver that spends several hours at work could park and charge their EV at a nearby station and walk from the station to work. For each station, we assume that drivers would choose to "walk" or "wait" to minimize the additional time spent refueling. Having specified the excess time to refuel at each charging station, we use our refueling choice model to predict drivers' EV refueling choices under different assumptions about both the speed of EV changing technology and the density of the EV charging network. We provide more details on the EV refueling excess time calculations in Appendix C.1.

We make several assumptions to simplify the counterfactual exercise. Firstly, we assume that all the EV stations have the same prices, charging speed, and brand quality,[29] and are accessible to all vehicles. In our baseline counterfactuals, we further assume that electric vehicles have the same fuel economy and range as the 2010 Honda Accord that drivers used in the IVBSS experiment and that the estimated relationship between "tank level" and value

[28]The distribution of the time to the closest electric charging station is right skewed because some trips are located very far from charging stations.

[29]Our estimated utility function includes station brand effects. For the counterfactuals, we omit these fixed effects and solve for a new intercept in the utility function associated with stopping to refuel such that the number of predicted EV refueling stops in the counterfactual equals the number of gasoline refueling stops observed in the data. See Appendix Section C.2 for more details.

of stopping to refuel is held fixed. Finally, we assume that drivers would not change their travel routes or schedules if they were to switch to an EV.

Table 6: Refueling Times by Technology Type

Type	Fill Rate	Technological Fill Time (Min.)	Excess Time Per Refuel: Total (Min.)	Drive (Min.)	Walk (Min.)	Wait (Min.)	Walk Share [0, 1]	Time Cost Per Stop ($)	Total Time Cost ($)
Gas									
Pump	10 Gal./Min.	0.77	2.54	1.78	0.00	0.76	0.00	1.17	829
Electric									
Average Charger Speed (2021)	29 kW	528.99	30.63	-0.12	30.10	0.64	1.00	14.06	9169
Tesla DC Supercharger	250 kW	61.36	18.26	-0.04	16.91	1.40	1.00	8.38	5131

Notes: For the "Gas Pump" technology type, we assume drivers consider only previously passed gasoline stations. For the "Electric" types, we assume that drivers consider all electric charging stations in the DOE database and that all the stations feature same charging technology and uniform prices. Waiting time for gas stations are based on an EPA rule that limits gas pumping speed to 10 gallons per minute. The technological fill time column indicates the amount of minutes needed to refuel based on the average refueling quantity in our estimation sample (7.65 gallons, equivalent to 255.59 KW). The excess time columns indicate the average amount of added time to refuel based on the station locations, refueling technology, and the driver preference estimates. The walk share column indicates that the fraction of refueling stops where it is time-minimizing for drivers to park at the charging station and then walk to their final destination. The average time cost per stop is calculated based on the baseline VOT estimate from the previous section. The total time cost is the the discounted expected time cost aggregated over the lifetime of a vehicle.

We discuss the implications and validity of the first two assumptions in Appendix C.3. While these assumptions may seem relatively strong, we are able to assess the sensitivity of our results to changes in these assumptions, which we discuss further below. The last assumption, that drivers would not change their routes excludes the possibility that driver re-optimize their driving trips to better accommodate EV charging. For example, a driver might choose to visit a different grocery store or cafe that has an EV charger nearby. This last assumption is difficult to relax without fully specifying a model of drivers' destination choices, which is outside the scope of this paper. Therefore, our EV refueling time cost estimates should be interpreted cautiously, with the caveat that drivers cannot change their trip destinations in the counterfactuals.

Our model suggests that EV drivers without home charging spend substantially more time on refueling than gasoline vehicle drivers. As shown in Table 6, in 2021, the average charging stop adds 30.6 excess minutes of driving, walking, and/or waiting time relative to only 2.5 excess minutes for a gasoline stop. Over the course of a vehicle's lifetime, our VOT estimates imply that excess refueling time costs for EVs are $9,169 compared to $829 for gas vehicles. The high time cost of EV refueling is driven by the slow speed of EV chargers. We estimate that the average charger speed in 2021 was 29 kW,[30] which means that it would take over 500 minutes to recharge 256 kWh of electricity (equivalent to 7.65 gallons of gasoline). Consequently, Table

[30] DOE does not report the exact speed of each charger so we calculate the average charging speed of each station under the assumption that chargers classified as Type 1, Type 2, or Type 3 (DC Fast) offer speeds of 1.92 kW, 19 kW, and 80 kW, respectively (SAE, 2017).

6 shows that drivers would typically spend nearly all of the 30.6 excess refueling minutes walking round-trip from the charging station to their destination.[31]

Although EV refueling time costs are still large compared gasoline, Table 5 shows that excess refueling times for EVs fell substantially in the decade between 2011 and 2021. Over that period, excess refueling time fell by nearly 75%, from 121.6 minutes to 30.6 minutes per refueling stop, due to the growing number of stations and the availability of fast chargers. The bottom row of Table 6 highlights the value of charging speed in reducing excess refueling time. If all electric chargers in 2021 were Tesla DC Superchargers (250 kW) rather than the observed average charging speed of 29 kW, drivers' excess time per stop would nearly halve to 18.2 minutes from 30.6 minutes. However, DC fast chargers are substantially more expensive than slower chargers: a 250 kW DC supercharger would cost approximately 10 times as much as a 25 KW charger (Nicholas, 2019). This raises an important policy question of whether investing in additional charging stations is more valuable than investing in increased charging station speed (with fewer stations).

We cast the charging network investment problem through the lens of a social planner who seeks a charging network that minimizes the refueling time cost of a representative driver subject to a budget constraint.[32] Namely, the planner chooses the optimal (second-best) combination of charging speed and number of stations.[33] The first order conditions of planner's problem—which we include in Appendix C.4—reveal an intuitive condition for optimality, the elasticity of excess refueling time with respect to the number of stations must equal the elasticity of excess refueling time ould with respect to charging speed.

To determine the relevant elasticities for the planners' investment problem, we solve our EV refueling choice model separately across a grid of combinations of station density (as observed between 2011 and 2021) and currently technologically feasible charging speeds (from 20 KW to 260 kW).[34] Figure 4 displays contour lines of excess refueling time per stop as well as the observed combinations of station density and charging speed over time (blue line). The figure makes clear that investment in electric charging stations has predominantly been to increase the number of charging stations rather than charging speed.

The first order conditions of the social planner's optimization indicate that the elasticity of excess refueling time with respect to charging speed should be set equal to the excess time

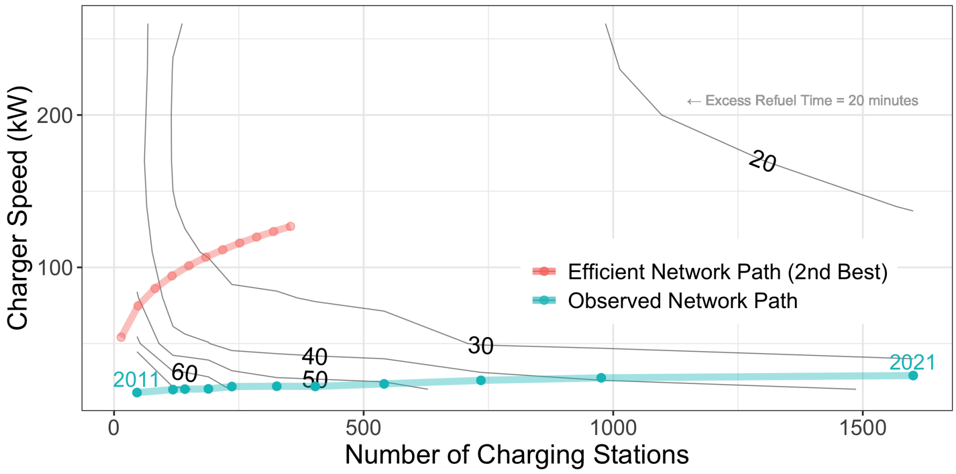
[31] If the value of time spent walking is different from the value of time spent driving to refuel, then this would change the valuation of this excess time.

[32] Casting the planner's problem with respect to a single representative driver increases the tractability of the problem, but abstracts away from explicitly modelling charging capacity constraints. Ignoring capacity constraints is perhaps more reasonable in this context because the choice between fewer fast chargers and a higher quantity of slow chargers will not necessarily alter the total charging capacity of the network. For example, one 100 kW fast charger could charger 10 vehicles back-to-back and achieve the same result as ten vehicles charging simultaneously at ten separate 10 kW AC chargers.

[33] We do not argue that this optimization is welfare maximizing, but rather that it is the most efficient way to spend a pre-specified charging infrastructure budget. Given that electric vehicle adoption depends on charging cost and vice-versa, the socially optimal level of charging station investment is a substantially more difficult question. See Greaker and Heggedal (2010) and Li et al. (2016) for a more in-depth discussion of these issues.

[34] See Appendix C.5 for more details.

Figure 4: Excess Refuel Time Contour Map as a Function of Station Density and Charger Speed



Notes: The thin grey lines show contours representing the estimated excess refueling time per EV refueling stop across different counterfactual combinations of station density (number of stations) and charging speed of the network (kW). The excess refueling time for each counterfactual network configuration is determined by the location of charging stations relative to drivers routes and behavioral assumptions that are described in Section 4. The lower blue line plots the evolution of the observed EV charging network density and speed from 2011 to 2021. The red line shows combination of charger density and speed that would minimize drivers' excess refueling time, while holding the total capital cost of the network fixed.

elasticity with respect to the number of stations. The red line in Figure 4 shows that the social planner's solution would be to invest in substantially fewer, but faster, chargers than was observed in the data. By 2021, the planner could have decreased EV refueling time costs by an additional 17% (30.6 minutes to 25.47 minutes per stop) by investing in faster chargers but fewer total stations than the observed network configuration.

We find that the marginal value of faster charging speeds has exceeded the marginal value of additional stations consistently over the past decade. Table A.11 illustrates that in 2011, the excess time elasticity with respect to the number of stations was -0.38, whereas the elasticity with respect to charging speed was -0.69. This suggests that the return to additional investment in charging speed was 1.76 times larger (-0.71/-0.39) greater than the return to investment in a proportional increase in the number of stations. By 2021, the time-savings from higher charging speed grew to be 4.72 times (-0.52/-0.11) greater than the time-savings from a proportional increase in the number of stations.

Overall, this combination of results shows that the marginal value of charging speed is substantially higher than the marginal value of charging station density, and that this disparity has been increasing over time. Hence, the counterfactuals suggest that policies that prioritize investments in faster chargers as opposed to additional chargers (with slower charging speeds), may deliver larger benefits to EV drivers.

Notably, the EV refueling results in this section rely on a number of strong assumptions. Appendix C.6 presents a collection of sensitivity checks. More specifically, we consider the implications of varying the assumed frequency that drivers would recharge their EV and altering how EV drivers' value waiting time relative to walking time. These sensitivity checks demonstrate that our excess refueling time estimates are somewhat sensitive to some of the underlying assumptions. For example, the estimated refueling times substantially increase from our baseline if drivers were to increase the frequency that they refuel their EVs. On the other hand, the sensitivity checks also show that the major takeaway that the marginal value of charging speed exceeds the marginal value of additional stations, remains unchanged under different assumptions about drivers' behavior and preferences.

6 Conclusion

While the economics literature has made substantial strides in understanding the supply and demand for gasoline, our ability to understand drivers' refueling choices has been limited by the lack of micro data on driver behavior. Understanding drivers' preferences—including their valuation of travel time—is important for both the retail gasoline market itself and for transportation policy more broadly. In particular, policies aiming to electrify the transportation sector will require a detailed understanding of the trade-offs drivers face when making refueling decisions.

In this paper, we use high frequency GPS driving data to better understand drivers' behavior and preferences for refueling. We leverage these unique data to document new facts about drivers' refueling choices–such as the distance driven out of the way to refuel. Moreover, we develop a model of refueling choice to recover estimates of drivers' preferences. We show that drivers rely substantially on long-run average prices when forming expectations of the prices they will pay at each station and that drivers have a high value of time. This high value of time suggests that current Department of Transportation benefit-cost analyses may under-value the benefits of time-saving transportation policies, and that policies mandating station price disclosure may have limited welfare benefits to drivers.

Drivers' high value of time has particularly important implications for EV policies. We use our model to measure the refueling costs and benefits drivers would receive if they drove EVs instead of gasoline vehicles. We find that drivers who charge at home benefit from reduced refueling time, while drivers who rely upon the public charging network face substantial additional costs. Drivers with home charging are generally wealthier, suggesting that policies to improve public charging infrastructure could make EV adoption attractive to a wider range of drivers. Finally, our results demonstrate that policies that prioritize faster charging stations over increased station density will generate larger decreases in refueling times for drivers who rely on public charging.

This paper contributes to the literature by providing direct evidence on drivers' on-road refueling decisions, but there is substantial room for future research to extend and potentially validate our results. In particular, our data come from a relatively small sample of drivers, and further analysis with a larger or more nationally representative sample would be valuable.

Additionally, as EVs become increasingly common, it will be important for researchers to empirically document drivers' choices when making charging decisions.

Beyond driver preferences, designing effective EV policy will require a more detailed understanding of the incentives facing EV charging stations. The EV charging market has thus far developed differently from the gasoline station market, with large networks of chargers posting fixed prices. Understanding why these differences have arisen and what they imply for the role of policy will be critical for encouraging widespread EV adoption.

References

Abaluck, J. and Adams-Prassl, A. (2021). What do consumers consider before they choose? identification from asymmetric demand responses. *The Quarterly Journal of Economics*, 136(3):1611–1663.

Allcott, H. (2013). The welfare effects of misperceived product costs: Data and calibrations from the automobile market. *American Economic Journal: Economic Policy*, 5(3):30–66.

Allcott, H. and Knittel, C. (2019). Are consumers poorly informed about fuel economy? evidence from two experiments. *American Economic Journal: Economic Policy*, 11(1):1–37.

Allcott, H. and Taubinsky, D. (2015). Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. *American Economic Review*, 105(8):2501–38.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal*, 75(299):493–517.

Beesley, M. E. (1965). The value of time spent in travelling: some new evidence. *Economica*, 32(126):174–185.

Borenstein, S., Cameron, A. C., and Gilbert, R. (1997). Do Gasoline Prices Respond Asymmetrically to Crude Oil Price Changes? *The Quarterly Journal of Economics*, 112(February):305–339.

Buchholz, N., Doval, L., Kastl, J., Matějka, F., and Salz, T. (2020). The value of time: Evidence from auctioned cab rides. Technical report, National Bureau of Economic Research.

Byrne, D. P. and de Roos, N. (2017). Consumer search in retail gasoline markets. *The Journal of Industrial Economics*, 65(1):183–193.

Byrne, D. P. and De Roos, N. (2019). Learning to coordinate: A study in retail gasoline. *American Economic Review*, 109(2):591–619.

Byrne, D. P., Leslie, G. W., and Ware, R. (2015). How do consumers respond to gasoline price cycles? *The Energy Journal*, pages 115–147.

Cardell, N. S. (1997). Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. *Econometric Theory*, 13(2):185–213.

Cesario, F. J. (1976). Value of time in recreation benefit studies. *Land Economics*, 52(1):32–41.

Chandra, A. and Tappata, M. (2011). Consumer search and dynamic price dispersion: an application to gasoline markets. *The RAND Journal of Economics*, 42(4):681–704.

Davis, L. W. (2019). Evidence of a homeowner-renter gap for electric vehicles. *Applied Economics Letters*, 26(11):927–932.

De los Santos, B., Hortaçsu, A., and Wildenbeest, M. R. (2012). Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review*, 102(6):2955–80.

Deacon, R. T. and Sonstelie, J. (1985). Rationing by waiting and the value of time: Results from a natural experiment. *Journal of Political Economy*, 93(4):627–647.

DOE (2022). Department of energy alternative fuels data center. https://afdc.energy.gov, note = Online; accessed 26 January 2022.

Ge, Y., Simeone, C., Duvall, A., and Wood, E. (2021). There's no place like home: Residential parking, electrical access, and implications for the future of electric vehicle charging infrastructure. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

Goeree, M. S. (2008). Limited Information and Advertising in the U.S. Personal Computer Industry. *Econometrica*, 76(5):1017–1074.

Goldszmidt, A., List, J. A., Metcalfe, R. D., Muir, I., Smith, V. K., and Wang, J. (2020). The Value of Time in the United States: Estimates from Nationwide Natural Field Experiments. Technical report, National Bureau of Economic Research.

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

Hastings, J. S. and Shapiro, J. M. (2013). Fungibility and consumer choice: Evidence from commodity price shocks. *The Quarterly Journal of Economics*, 128(4):1449–1498.

Hotz, V. J. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60(3):497–529.

Houde, J.-F. (2012). Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline. *American Economic Review*, 102(5):2147–2182.

Houde, S. (2018). How consumers respond to product certification and the value of energy information. *The RAND Journal of Economics*, 49(2):453–477.

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

Ito, K. (2014). Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing. *American Economic Review*, 104(2):537–63.

Ito, K. and Sallee, J. M. (2018). The economics of attribute-based regulation: Theory and evidence from fuel economy standards. *Review of Economics and Statistics*, 100(2):319–336.

Knittel, C. R. and Tanaka, S. (2021). Fuel economy and the price of gasoline: Evidence from fueling-level micro data. *Journal of Public Economics*, 202:104496.

Lave, C. A. (1969). A behavioral approach to modal split forecasting. *Transportation Research/UK/*, 3(4).

Leggett, C. G. (2002). Environmental valuation with imperfect information the case of the random utility model. *Environmental and Resource Economics*, 23(3):343–355.

Levin, L., Lewis, M. S., and Wolak, F. A. (2017). High frequency evidence on the demand for gasoline. *American Economic Journal: Economic Policy*, 9(3):314–47.

Levin, L., Lewis, M. S., and Wolak, F. A. (2019). Reference dependence in demand for gasoline.

Lewis, M. S. (2011). Asymmetric Price Adjustment and Consumer Search: An Examination of the Retail Gasoline Market. *Journal of Economics and Management Strategy*, 20(2):409–449.

Lewis, M. S. and Marvel, H. P. (2011). When Do Consumers Search? *The Journal of Industrial Economics*, LIX(3):457–483.

Lewis, M. S. and Noel, M. D. (2011). The Speed of Gasoline Price Response in Markets with and without Edgeworth Cycles. *Review of Economics and Statistics*, 93(May):672–682.

Li, J. (2017). Compatibility and investment in the U.S. electric vehicle market. Technical report.

Li, S., Lang, T., Xing, J., and Zhou, Y. (2016). The Market for Electric Vehicles: Indirect Network Effects and Policy Design. *Journal of the Association of Environmental and Resource Economists*, 4(1):89–133.

Li, Jing (2017). Compatibility and Investment in the U.S. Electric Vehicle Market.

Liebman, J. B. and Zeckhauser, R. J. (2004). Schmeduling. Technical report.

Lu, S. (2006). Vehicle survivability and travel mileage schedules. Technical report.

Luco, F. (2019). Who benefits from information disclosure? The case of retail gasoline. *American Economic Journal: Microeconomics*, 11(2):277–305.

Luxen, D. and Vetter, C. (2011). Real-time routing with openstreetmap data. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, GIS '11, pages 513–516, New York, NY, USA. ACM.

Nicholas, M. (2019). Estimating electric vehicle charging infrastructure costs across major us metropolitan areas. *URL: https://theicct. org/sites/default/files/publications/ICCT_EV_Charging_Cost_20190813. pdf.*

Oort, C. J. (1969). The evaluation of travelling time. *Journal of Transport Economics and Policy*, pages 279–286.

Pennerstorfer, D., Schmidt-Dengler, P., Schutz, N., Weiss, C., and Yontcheva, B. (2020). Information and price dispersion: Theory and evidence. *International Economic Review*, 61(2):871–899.

SAE (2017). *Society of Automotive Engineers Electric Vehicle and Plug in Hybrid Electric Vehicle Conductive Charge Coupler.*

Schmeiser, S. (2014). Consumer inference and the regulation of consumer information. *International Journal of Industrial Organization*, 37:192–200.

Small, K. A. and Rosen, H. S. (1981). Applied welfare economics with discrete choice models. *Econometrica: Journal of the Econometric Society*, pages 105–130.

Small, K. A., Winston, C., and Yan, J. (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, 73(4):1367–1382.

Springel, K. (2016). Network externality and subsidy structure in two-sided markets: Evidence from electric vehicle incentives. *Job Market Paper*, 2016.

Teisberg, T. J. (1981). A dynamic programming model of the us strategic petroleum reserve. *The Bell Journal of Economics*, pages 526–546.

Train, K. (2015). Welfare calculations in discrete choice models when anticipated and experienced attributes differ: A guide with examples. *Journal of Choice Modelling*, 16:15–22.

Traut, E. J., Cherng, T. C., Hendrickson, C., and Michalek, J. J. (2013). U.S. residential charging potential for electric vehicles. *Transportation Research Part D: Transport and Environment*, 25:139–145.

U.S. Bureau of Labor Statistics (2020). Consumer Expenditure Survey. Technical report, U.S. Bureau of Labor Statistics.

White, V. (2016). Revised departmental guidance on valuation of travel time in economic analysis. *Office of the Secretary of Transportation, US Department of Transportation, Available at: transportation.gov/sites/dot. gov/files/docs/2016%20Revised%20Value%20of%20Travel%20Time%20Guidance. pdf.*

Wolff, H. (2014). Value of time: Speeding behavior and gasoline prices. *Journal of Environmental Economics and Management*, 67(1):71–88.

Yang, H. and Ye, L. (2008). Search with Learning: Understanding Asymmetric Price Adjustments. *The RAND Journal of Economics*, 39(2):547–564.

Online Appendix

A Additional Tables & Figures Referenced in Main Paper

Table A.1: Driver Summary Statistics

	Mean	SD	Min	Max
Driver Census Tract Median Income ($)	64,275.44	26,074.66	19,710.00	146,250.00
Days With Vehicle	39.83	4.18	32.00	73.00
Total Driving Distance (Miles)	1,768.48	862.34	520.98	4,585.83
Miles Per Day	50.03	22.34	13.45	127.38
Total Driving Trips	204.76	90.33	37.00	597.00
Total Number of Refueling Stops	8.01	5.49	0.00	32.00
Refueling Stops Per Week	1.57	0.99	0.00	6.22
Observations	108			

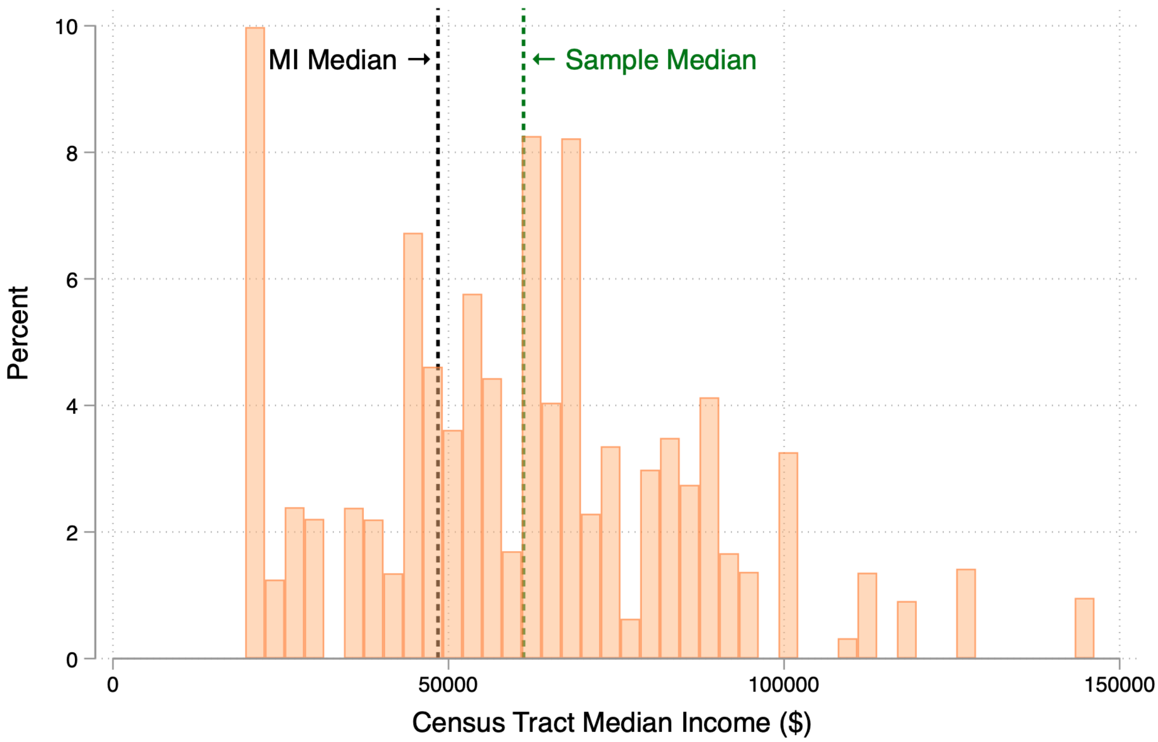
Notes: Panel A summarizes information about the drivers including their driving behavior during the IVBSS experiment. The drivers were 50% male and 50% female. The drivers were approximately evenly distributed across three age groups: 20-30 years old, 40-50 years old, and 60-70 years old.

Table A.2: Summary of Vehicle Trips

	Mean	SD	Pct25	Median	Pct75
Trip Distance (miles)	8.27	16.45	1.04	3.50	9.00
Trip Time (minutes)	13.48	20.35	3.11	7.98	16.69
Distance from Trip Origin to Home (miles)	9.48	25.73	0.09	2.53	8.57
Weekend (0,1)	0.26	0.44	0.00	0.00	1.00
Refueling Stop (0,1)	0.04	0.19	0.00	0.00	0.00
Tank Level at Start of Trip (gallons)	6.87	4.21	3.74	6.55	9.77
Number of Gas Stations Available Nearby Optimal Route					
Within 1 minute(s)	5.89	6.47	2.00	4.00	7.00
Within 5 minute(s)	24.45	36.07	5.00	12.00	27.00
Within 20 minute(s)	120.45	149.08	36.00	71.00	142.00
Observations	22114				

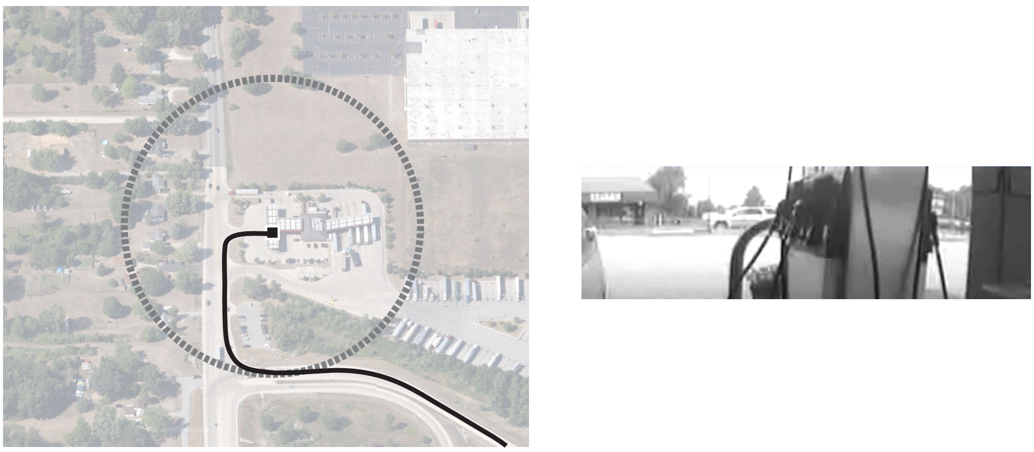
Notes: Summary statistics are reported across all trips made by all drivers during the experiment. Pct25 and Pct75 are the 25th and 75th percentiles, respectively.

Figure A.1: Driver Income Distribution



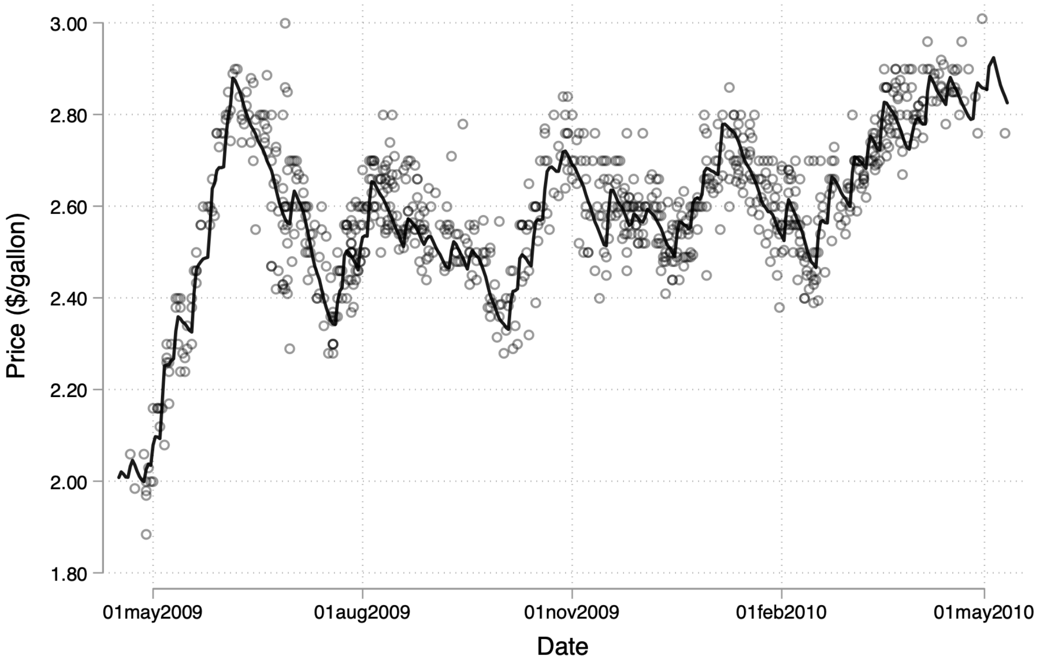
Notes: Distribution of driver's income based on the census tract in the 2010 American Community Survey.

Figure A.2: Procedure for identifying gas station stops



Notes: 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 A.3: Average gasoline price and observed purchase price through sample period



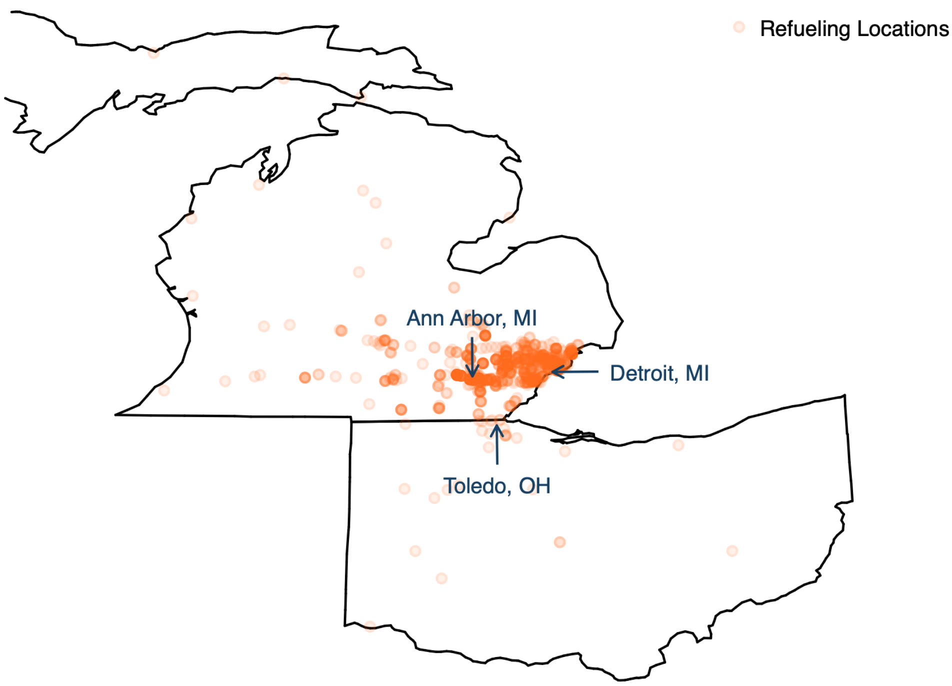
Notes: The black line shows the daily average gasoline price in MI and OH from the OPIS data. Each small dot represents the date and price associated with one of the refueling stops that we identify.

Table A.3: Station Brand Choice Probabilities

BP	19.7%
Citgo	5.0%
Costco	2.5%
Marathon	14.1%
Meijer	7.4%
Mobil	10.2%
Other	13.5%
Shell	6.4%
Speedway	14.8%
Sunoco	6.5%
Total	100.0%

Notes: Choice probabilities are shares across all refueling stops made by all drivers during the experiment.

Figure A.4: Map of Observed Refueling Locations



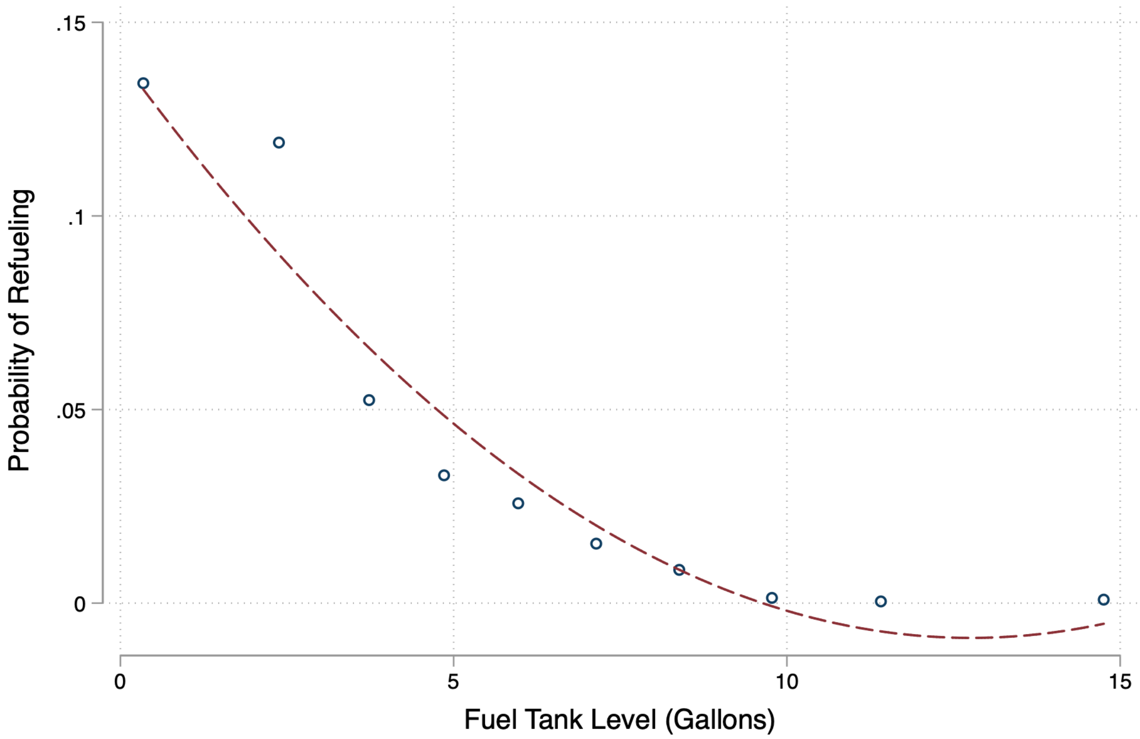
Notes: The orange points represent a refueling stop that we identified using vehicle geolocations and the in-car cameras.

Table A.4: Passed vs. Non-Passed Stations

	Passed		Non-Passed		Difference	
	Mean	SD	Mean	SD	Diff.	SE
Current Price ($/gallon)	2.620	0.140	2.589	0.165	-0.031***	(0.001)
Excess Time (min)	8.295	5.818	10.705	5.298	2.410***	(0.032)

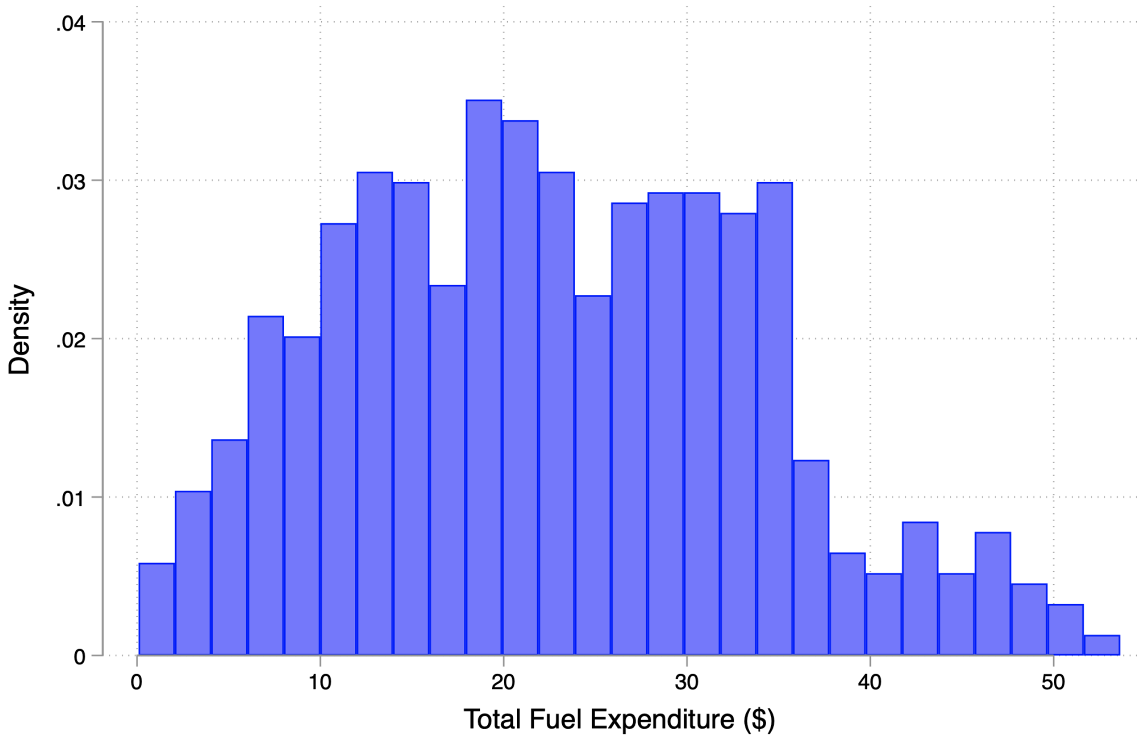
Notes: The table shows prices and excess time for all stations that were passed previously by drivers compared to stations that were not passed previously by drivers. The sample includes all potential stations that are within 20 minutes deviation from a trip.

Figure A.5: Probability of Stopping and Tank Level



Notes: Each dot in the binned scatter plot show how the empirical probability of stopping to refuel changes as function of fuel tank level measured at the start of the trip. The dashed line shows a quadratic fit of the data.

Figure A.6: Distribution of Total Fuel Expenditure



Notes: The graphic shows a histogram of total fuel expenditure in dollars across all refueling stops in our sample. Total expenditure is calculated as the current price ($ per gallon) at the driver's selected station multiplied by the implied fill quantity (gallons) for the trip. Our procedure for recovering fill quantities is described in the appendix.

Table A.5: Fill Quantity Regression - Evidence for Exclusion of Prices

	Fill Quantity (gallons)		
	(1)	(2)	(3)
(Intercept)	8.854	11.75	1.755
	(0.4268)	(2.938)	(9.665)
Tank Level	-0.4226	-0.4314	-0.4111
	(0.2148)	(0.2150)	(0.2157)
(Tank Level)^2	-0.0125	-0.0122	-0.0137
	(0.0275)	(0.0275)	(0.0275)
Current Price		-1.104	-1.358
		(1.110)	(1.134)
Station's Average Price			4.129
			(3.805)
Observations	744	744	744
R^2	0.10267	0.10389	0.10533
Station Brand Fixed Effects	Y	Y	Y
Choice Set	Passed	Passed	Passed

Notes: The dependent variable is the imputed fill quantity associated with each observed refueling stop. The regressions estimates are used to predict expected fill quantity conditional on stopping for each trip conditional on initial tank level (gallons). The first regression is used to predict expected fill quantity for the baseline model (6) in Table 3. The second two specifications provide evidence for the exclusion restriction imposed by Assumption 1. Neither current price nor station average price are statistically significant predictors of fill quantity after controlling for tank level and station brand.

Table A.6: Heterogeneous Preferences

	(1)	(2)	(3)	(4)	(5)
Expenditure Coefficients					
α (constant)	-0.246	-0.371	-0.588	-0.395	-0.384
	(0.042)	(0.040)	(0.306)	(0.041)	(0.174)
$\times$ 1[Age 60-70]	-0.552				-0.235
	(0.080)				(0.079)
$\times$ 1[Female]		-0.125			-0.085
		(0.054)			(0.061)
$\times$ 1[Low Income (Q1)]			0.293		0.030
			(0.423)		(0.144)
$\times$ 1[High Income (Q5)]			0.260		0.175
			(0.323)		(0.145)
$\times$ 1[Weekend]				-0.055	0.054
				(0.058)	(0.033)
Weight on Current Price					
θ (constant)	0.177	0.241	0.104	0.176	0.187
	(0.144)	(0.062)	(0.042)	(0.064)	(0.146)
$\times$ 1[Age 60-70]	-0.084				-0.188
	(0.150)				(0.145)
$\times$ 1[Female]		-0.197			-0.105
		(0.084)			(0.228)
$\times$ 1[Low Income (Q1)]			0.111		0.106
			(0.420)		(0.190)
$\times$ 1[High Income (Q5)]			0.386		0.350
			(0.323)		(0.348)
$\times$ 1[Weekend]				-0.015	0.034
				(0.111)	(0.188)
Excess Time					
γ	-0.152	-0.162	-0.175	-0.159	-0.156
	(0.013)	(0.013)	(0.038)	(0.015)	(0.116)
Nesting Parameter					
λ	0.297	0.321	0.350	0.316	0.309
	(0.024)	(0.027)	(0.086)	(0.027)	(0.141)
Number of Stops	744	744	744	744	744
Number of Trips	21355	21355	21355	21355	21355
Observations	815070	815070	815070	815070	815070
Choice Set	Passed	Passed	Passed	Passed	Passed

Notes: Table reports psuedo maximum likelihood estimates of driver preferences. Bootstrap standard errors are reported in parentheses. The choice set for each specification includes stations that the driver has previously passed before that are within 20 minutes of the route.

Table A.7: Average Marginal Effects ($/Hour)

1[Age 60-70]	1[Female]	1[Low Income (Q1)]	1[High Income (Q5)]	1[Weekend]
-11.18	-5.23	2.14	17.09	3.91
(7.45)	(6.54)	(5.52)	(8.24)	(5.28)

Notes: Table reports the average marginal effect on the value of time. The marginal effects corresponds to the estimates in Column 5 of Table A.6. Standard errors are reported in parentheses.

Table A.8: Driver Preference Estimates - Varying Specification of Driver Choice Sets

	(1)	(2)	(3)	(4)	(5)
Decision to Stop					
1[Stop] $\times$ Constant	6.414	6.019	4.948	7.110	8.174
	(1.193)	(1.354)	(1.472)	(1.276)	(1.371)
1[Stop] $\times$ Tank Level	-0.453	-0.376	-0.367	-0.477	-0.534
	(0.075)	(0.084)	(0.091)	(0.077)	(0.082)
1[Stop] $\times$ (Tank Level)^2	-0.060	-0.066	-0.062	-0.063	-0.060
	(0.007)	(0.009)	(0.010)	(0.007)	(0.007)
Station Choice					
α - Expenditure	-0.384	-0.411	-0.362	-0.414	-0.462
	(0.053)	(0.067)	(0.074)	(0.056)	(0.061)
θ - Weight on Current Price	0.169	0.156	0.245	0.170	0.145
	(0.068)	(0.077)	(0.101)	(0.066)	(0.060)
γ - Excess Time (minutes)	-0.176	-0.164	-0.172	-0.183	-0.181
	(0.017)	(0.019)	(0.022)	(0.018)	(0.018)
Nesting Parameter					
λ	0.350	0.333	0.326	0.368	0.371
	(0.033)	(0.038)	(0.042)	(0.034)	(0.035)
Station Brand Fixed Effects	Y	Y	Y	Y	Y
Number of Stops	744	549	441	723	692
Number of Trips	21355	17158	13960	21290	21125
Observations	815070	675126	527458	694467	553320
Choice Set	Passed Ever	Home $\leq$ 10 mi. & Passed	Home $\leq$ 5 mi. & Passed	Passed $\leq$ 14 days	Passed $\leq$ 7 days
	Implied Value of Time ($/hour)				
	27.54	23.98	28.40	26.47	23.53
	(3.31)	(3.28)	(4.95)	(3.19)	(2.70)

Notes: Table reports pseudo maximum likelihood estimates of driver preferences. The expected fill quantities are predicted from the regressions in Table 2. Each column shows results varying our specification of drivers' choice set. Column (1) shows our base specification where all stations that the driver has previously passed that are within 20 minutes of the route are included in the choice set. Column (2) and (3) also set the choice set to stations that the driver has previously passed that are within 20 minutes, but further restrict the sample to only trips that started within 10 miles and 5 miles of the driver's home, respectively. Column (4) and (5) restrict the choice sets to only stations within 20 minutes of the route and that the driver has passed within the last 14 days or 7 days, respectively. Coefficient standard errors are reported in parentheses. The implied value of time is calculated as $60 \cdot \frac{\gamma}{\alpha}$ and standard errors are reported in parentheses.

Table A.9: Driver Preference Estimates - Varying Specification of $\overline{P}$

	(1)	(2)	(3)	(4)	(5)
Decision to Stop					
1[Stop] $\times$ Constant	6.414	0.844	1.386	3.546	3.096
	(1.193)	(0.515)	(0.600)	(0.737)	(0.730)
1[Stop] $\times$ Tank Level	-0.453	-0.186	-0.212	-0.317	-0.287
	(0.075)	(0.054)	(0.056)	(0.060)	(0.059)
1[Stop] $\times$ (Tank Level)^2	-0.060	-0.052	-0.053	-0.056	-0.056
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Station Choice					
α - Expenditure	-0.384	-0.141	-0.164	-0.258	-0.238
	(0.053)	(0.022)	(0.026)	(0.033)	(0.032)
θ - Weight on Current Price	0.169	0.855	0.461	0.155	0.273
	(0.068)	(0.186)	(0.251)	(0.103)	(0.108)
γ - Excess Time (minutes)	-0.176	-0.168	-0.164	-0.163	-0.164
	(0.017)	(0.018)	(0.018)	(0.017)	(0.017)
Nesting Parameter					
λ	0.350	0.336	0.327	0.323	0.326
	(0.033)	(0.035)	(0.035)	(0.033)	(0.034)
Station Brand Fixed Effects	Y	Y	Y	Y	Y
Number of Stops	744	744	744	744	744
Number of Trips	21355	21355	21355	21355	21355
Observations	815070	815070	815070	815070	815070
Choice Set	Passed	Passed	Passed	Passed	Passed
$\overline{P}$	Full Sample	Last Price Passed	Month	Quarter	Half-Year
Log Likelihood	-4374.128	-4394.963	-4393.544	-4379.452	-4384.249
	Implied Value of Time ($/hour)				
	27.54	71.64	60.27	37.84	41.20
	(3.31)	(12.29)	(10.60)	(5.07)	(5.92)

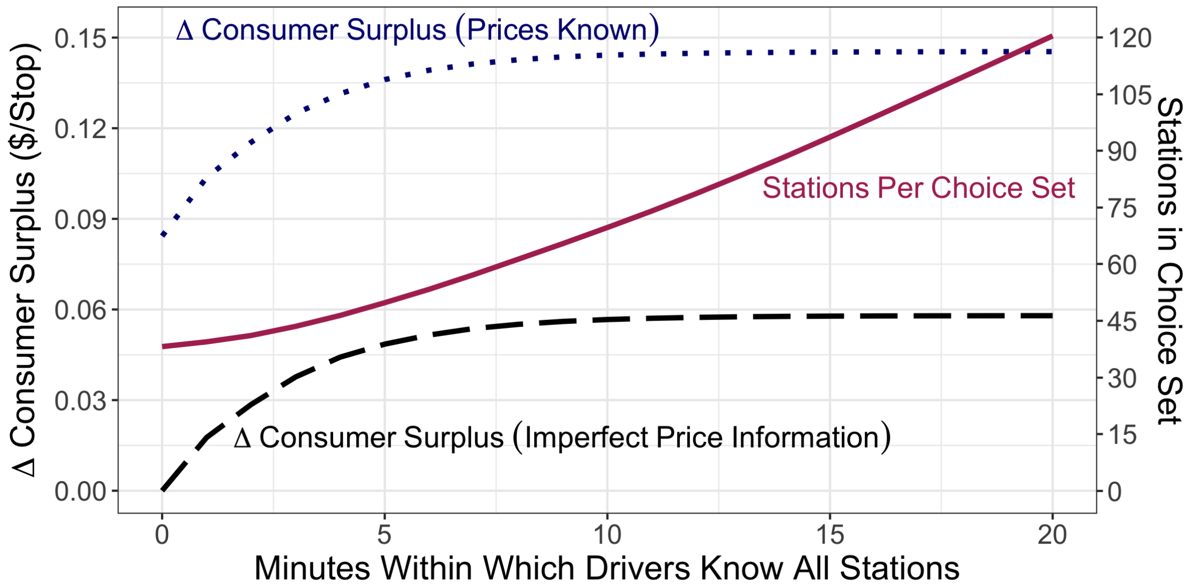
Notes: Table reports pseudo maximum likelihood estimates of driver preferences. The expected fill quantities are predicted from the regressions in Table 2. Each column shows how estimates change if we specify a different measure of average price ($\overline{P}$) entering the drivers expectations. The first column is our base specification that sets $\overline{P}$ as the station's average price over the entire sample, column (2) sets $\overline{P}$ as the price the last time the driver passed the station, columns (3-5) set $\overline{P}$ as the mean price at the station over the current month, quarter, and half year, respectively. Choice Set = "All" indicates that all stations with 20 minutes of the driver's route are included in the choice set. Choice Set = "Passed" means stations that the driver has previously passed that are within 20 minutes of the route are included in the choice set. Coefficient standard errors are reported in parentheses. The implied value of time is calculated as $60 \cdot \frac{\gamma}{\alpha}$ and standard errors are reported in parentheses.

Table A.10: Driver Preference Estimates - Station Neighborhood Controls

	(1)	(2)	(3)	(4)
Decision to Stop				
1[Stop] $\times$ Constant	6.414	6.560	7.041	7.048
	(1.193)	(1.215)	(1.271)	(1.278)
1[Stop] $\times$ Tank Level	-0.453	-0.461	-0.479	-0.480
	(0.075)	(0.076)	(0.078)	(0.078)
1[Stop] $\times$ (Tank Level)^2	-0.060	-0.060	-0.061	-0.061
	(0.007)	(0.007)	(0.007)	(0.007)
Station Choice				
α - Expenditure	-0.384	-0.392	-0.410	-0.410
	(0.053)	(0.054)	(0.056)	(0.057)
θ - Weight on Current Price	0.169	0.166	0.162	0.161
	(0.068)	(0.067)	(0.066)	(0.066)
γ - Excess Time (minutes)	-0.176	-0.178	-0.187	-0.187
	(0.017)	(0.017)	(0.018)	(0.018)
Station Census Tract Median Income		0.004		0.000
		(0.008)		(0.008)
Station Census Tract Population Density			-0.035	-0.035
			(0.014)	(0.014)
Nesting Parameter				
λ	0.350	0.354	0.372	0.372
	(0.033)	(0.033)	(0.035)	(0.035)
Station Brand Fixed Effects	Y	Y	Y	Y
Number of Stops	744	744	744	744
Number of Trips	21355	21355	21355	21355
Observations	815070	815070	815070	815070
	Implied Value of Time ($/hour)			
	27.54	27.28	27.30	27.28
	(3.31)	(3.27)	(3.25)	(3.27)

Notes: Table reports pseudo maximum likelihood estimates of driver preferences. The expected fill quantities are predicted from the regressions in Table 2. Each column shows how estimates change if we add controls for the characteristics for the station's neighborhood. In particular, we add controls for the median household income ($10,000s) and population density (100 inhabitants per square mile) of each station's census tract according to data from the 2010 American Community Survey. Coefficient standard errors are reported in parentheses. The implied value of time is calculated as $60 \cdot \frac{\gamma}{\alpha}$ and standard errors are reported in parentheses.

Figure A.7: Consumer Surplus Gains from Adding Stations to Choice Set



Notes: The solid maroon line plots the average number of stations available to drivers if all the previously unpassed stations within X minutes of the drivers' optimal route were added to the choice set. The dashed black line shows the change in consumer surplus in dollars per gas stop from adding additional (unpassed) stations to the choice set, assuming that drivers remain imperfectly informed about current gas prices, relative to the baseline case where only passed stations are in the choice set. The dotted blue line shows the change in consumer surplus in dollars per gas stop if drivers were perfectly informed about current prices and additional unpassed stations are added to the choice set, relative to the baseline case with only passed stations in the choice set and imperfect information about current prices.

Table A.11: Driver Access to Gasoline Station Network vs. Electric Charging Station Network

	Gas (All)	Gas (Passed)	Electric (All)
Stations within 5 Minutes of Route	37.17	23.41	4.75
Closest Station (minutes)	0.77	0.98	4.31

Notes: The first column includes all gas stations in OPIS and the second column only counts the stations that each driver has previously passed. The third column summarizes all public electric chargers in 2021 based on DOE data. The first row reports the average number of stations located within five minutes deviation from a driver's route, across all trips in our data. The second row shows that average time to the closest station relative to driver's routes across all trips in the data.

B Recovering Fuel Tank Levels and Fill Quantities

Throughout the IVBSS experiment on-board computers collected an abundance of high frequency data. One variable that the on-board computers did not directly record is each vehicle's current fuel tank level. Fortunately, the computers did record each vehicle's instantaneous fuel consumption (*i.e.* the rate of fuel usage). In addition, each vehicle was equipped with an over-the-shoulder camera which allows us to observe photos of the dashboard fuel gauge. We combine the data on instantaneous fuel consumption together with the information from the fuel gauge photos to recover estimates of the current fuel tank level for each vehicles at the start of each trip.

Our procedure to recover the fuel tank levels involves two primary parts. In the first part, Amazon Mechanical Turk workers viewed images of the fuel gauge and classified each image into numeric tank level category. We then use the fuel tank classifications along with the data of instantaneous fuel consumption in a regression framework to recover implied fuel tank levels. This procedure also provides us an estimate of the quantity of fuel that each driver purchased at each refueling stop. The details of our procedure are as follows:

1. First, we collected the sample of every-five-minute cabin photos. From the cabin photos, we cropped a 30x25 pixel rectangle at the fuel gauge location. Based on an analysis of the pixels in the center of this rectangle, we classify the image into one of four types: good, underexposed, overexposed, and low contrast. The image is rescaled and smoothed.

2. We then uploaded the "good" gauge photos to Amazon S3 for classification by Mechanical Turk workers.

    - The Amazon Turk workers completed a series of tasks. Each task consisted of the classification of three fuel gauge photos. The tank levels ranged from 0 (empty) to 8 (full), with a value of 9 corresponding to an illegible tank level. At least three workers classified each photo and some photos were classified by six workers.

3. We estimate the tank level associated with each five minute interval (*i.e.* each photo) as the mean tank level classification (across the Amazon Turk workers) between 0 and 8. After using this data to impute the fuel tank levels, we rechecked and corrected some outlying classifications.

4. Next, we collected the raw IVBSS data for fuel consumption for each vehicle and trip. The fuel consumption variable records the cumulative fuel consumption (in milliliters) for each five-minute interval. Some corrections are required for this variable due to an overflow error on long trips, as well as some other errors caused by computer resets. We also collect the data on the exact times that refueling stops occurred. We then combine the incremental fuel consumption for each five-minute interval and the data on refueling times into a common data set.

5. To recover the fuel tank level at each five-minute interval for each driver, we estimate a regression of the cumulative fuel consumption on (i) a quadratic in the fuel gauge level between 0 and 8, and (ii) driver-by-refueling-event fixed effects.

   - The driver-by-refueling-event fixed effects correspond to the initial fuel tank level at the start of each refueling event. We also shift all of the tank levels up by two gallons based on an assumption that there are two gallons remaining in the fuel tank when the gauge is at empty.

6. We combine the initial fuel tank level at the time of each refueling event (from our estimated regressions), with the incremental fuel consumption data to calculate the current fuel tank level at each five-minute interval.[35]

7. Finally, we aggregate the data to the driver-by-trip level to obtain the initial fuel tank level at the start of each trip. In addition, we infer the fuel purchase quantity associated with each refueling event as the change in fuel tank level (in gallons) between trips before and after a refueling event occurred.

[35]We use the second-by-second fuel consumption data to obtain an estimate of the fuel tank level at each second and then aggregate up to the five-minute level.

C EV Refueling Choice Model and Counterfactuals

C.1 Determining Excess Refueling Times for EVs

When making their EV refueling decision, we allow drivers to either (1) drive to the charging station and wait for their vehicle to recharge or (2) park their vehicle at the charging station and walk to their destination. For each station, we assume that drivers would choose to "walk" or "wait" to minimize the excess time spent refueling.

We define excess time refueling as the additional time that a driver would spend if they choose to refuel on a given trip compared to if they were instead to travel directly from their trip's origin to destination on the optimal route plus the time spent at the destination. We include the time spent at the destination to incorporate that drivers can recharge their vehicle while they are visiting a destination. For example, a driver that spends several hours at work, could park and charge their EV at a station during the work day. Therefore, the calculation of excess time refuel time for each EV station on each trip entails several steps:

1. Determine the drivers' refuel quantities conditional on stopping.
    - In our estimation sample, drivers purchase an average of 7.65 gallons which is equivalent to 256 kWh of electricity. Thus, on our baseline counterfactual we assume that drivers would refill their EV with an equivalent amount of "fuel" as we observe the drivers refueling in the gasoline market. However, we also show how the results would change if drivers decided to refill their EV more frequently with smaller quantities
2. Calculate the technological time required for the driver to refuel.
    - The time required to refill is determined by the assumed speed of the charging technology. For example, with a 128 kW charger would take two hours to charge 256 kWh of energy.
3. Calculate the excess time associated associated with the drivers two possible refueling options: (1) "wait" and (2) "walk".
    - For the "wait" option the total excess refueling time is equal to the sum of the excess driving time to travel to the station plus the technological charging time (from Step 2). The excess drive time is calculated in the same way as we calculate excess travel time for gasoline stations (see Section 2).
    - For the "walk" method, excess time is calculated as follows:
        (a) Determine the amount of time that the driver spends at the final destination.
        (b) Calculate the time it would take the driver to walk round-trip from the refueling station to the destination assuming a walking speed of 3 miles per hour.

(c) Determine how much "additional" waiting time, if any, is needed to complete the charging cycle. Here, we compare the technological refuel time with the sum of the time spent at the destination and the round-trip walking time from the station. If the technological refuel time exceeds the sum, then additional waiting time is added to the total excess refueling time.
   - Suppose a driver requires three hours of technological charging time to achieve the refuel quantity established in Step 1. Further, suppose that the driver spends 2 hours at the destination and it takes 20 minutes to walk from the charging station to the destination. In this case, 40 minutes of waiting time is added to the total excess time for the "walk" method.

(d) Calculate the net driving time to drive from the origin to the charging station instead of the origin to the destination.
   - Note that this net driving time could be negative if the station is closer to the origin than the destination.

(e) The total excess time for the "walk" option is:
Total Excess Time = Added Walk Time (b) + Added Wait Time (c) + Net Drive Time (d)

4. The excess time for each station on each trip is determined by the option (walk or wait) with the minimum total excess time.

   - In our main specification we assume that drivers simply minimize time but we also show the robustness of the result under an alternative assumption that drivers strictly prefer waiting to walking and vice versa.

Example

As an example, suppose a driver is traveling from origin A to destination C and is considering stopping at station B. The driver has two options: (1) they can wait at station B—this option adds five minutes of driving time to visit station B, or (2) they can leave their car at station B and walk to destination C—this option saves two minutes of driving time but adds 32 minutes of round-trip walking time. The vehicle will take one hour to recharge and the driver plans to spend 20 minutes at destination C. If the driver waits at the station for the vehicle to charge, they will add 65 minutes of excess time to refuel—five minutes of added driving time plus 60 minutes of waiting time. On the other hand, if they choose to park and walk to the destination they would add only 40 minutes of excess refueling time—two fewer minutes of driving time, 30 minutes of added walking time, and 10 minutes of waiting time. Thus, we specify that the excess time associated with recharging at station B is 40 minutes.

C.2 EV Refueling Choice Model Implementation Details

After we calculate the excess refueling time for each EV station located within a 20-minute drive of each trip in our data we can utilize our refueling choice model outlined in Section 3 to predict which charging stations drivers would choose and the expected time cost associated with those refueling choices. To fully specify the utility model for our counterfactuals, we assume that all of EV stations charge the same prices and have the same brand quality. Our estimated utility function includes station brand effects, so for the counterfactuals, we omit these fixed effects and solve for a new intercept in the utility function associated with stopping to refuel such that the number of predicted EV refueling stops in the counterfactual equals the number of gasoline refueling stops observed in the data. In alternate versions of the counterfactuals, we assume that drivers stop more (or less) frequently to refuel but with smaller (or larger) quantities. For these alternate counterfactuals, we solve for the intercept in the utility function to match our assumed frequency of stops with the model's predicted stop frequency.

C.3 Discussion of Key EV Refueling Assumptions

1. Homogeneous prices and brand quality

    - We assume homogeneous prices and brand qualities in part because we lack detailed data on EV charging stations' prices and quality. However, this assumption also allows us to isolate the impact of changes in the density of charging stations and speed of the charging stations on excess refueling time. In practice, we would expect that charging stations with more convenient locations would be likely to charge higher prices. If stations with better locations were to charge higher prices that would lead us to underestimate expected EV refueling times. However, given that the EV network is relatively sparse and that drivers have a high value of time, we do not expect that price heterogeneity would lead to substantial changes in station choices.

2. EV range and frequency of refueling

    - In our main set of counterfactuals, we assume that EVs have the same range to gasoline vehicles and that drivers would choose to refuel them at similar frequencies and energy quantities. A benefit of this assumption is that it allows us to better isolate the effects of the charging station network on excess refueling time. However, a limitation of the assumptions is that current EVs may have shorter range than gasoline vehicles. Moreover, because EVs refuel at a much slower rate than gas vehicles, drivers may choose to refuel EVs more (or less) frequently with smaller (larger) quantities. We run several alternative specifications that vary the assumed frequency and quantities that drivers would refuel thier EV. We discuss the results implied by these assumptions in Section C.6.

3. Preference across refueling options: walking vs. waiting

   - In our main set of counterfactuals, we assume that drivers always choose the time-minimizing option whenever they decide to wait at a charging station or to park at the station and walk to their destination. In practice, drivers may have explicit preferences for either walking or waiting when recharging their EV. Therefore, we also solve the counterfactuals under two alternative assumptions: (1) drivers prefer to always walk and (2) drivers prefer to always wait. The results are in Section C.6.

C.4 Model of Public Charging Investment

Consider a planner that chooses a charging network design to minimizes drivers' time costs subject to a budget constraint. The planner can choose to build additional stations, N or to upgrade the speed, S, of each charger in the network. For simplicity, we assume that additional stations would be placed evenly apart in the network and that all chargers in the network have the same speed. Drivers' refueling time, τ is a decreasing function both the N and S. However, increasing either N or S will raise the capital cost of the network. Thus, the planner's investment problem can be written formally as:

$$\begin{aligned} \min_{N,S} \quad & \tau(N,S) \\ \text{s.t.} \quad & \kappa \cdot N \cdot S \leq \overline{B}. \end{aligned} \tag{C.1}$$

Here, $\overline{B}$ is the budget available to the planner to spend on charging infrastructure. The functional form for capital costs is motivated by Nicholas (2019) whose estimate show that capital cost are roughly proportional by the total power capacity $(N \cdot S)$ of the network. For example, installing three 50 kW chargers would cost approximately the same as installing one 150 kW charger. Therefore, the κ parameter represents the fixed cost of increasing the power capacity of the network.

Assuming the planner uses the entire budget, the following Lagrangian characterizes the solution of the planner's problem:

$$\mathcal{L}(N,S,\lambda) = \tau(N,S) - \lambda(\kappa \cdot N \cdot S - \overline{B}) \ . \tag{C.2}$$

FOCs

$$\frac{\partial \tau}{\partial S} = \lambda \cdot \kappa \cdot N \tag{C.3}$$

$$\frac{\partial \tau}{\partial N} = \lambda \cdot \kappa \cdot S \tag{C.4}$$

Rearranging the two FOCs we can derive the following simple optimality condition:

$$\tfrac{\partial \tau}{\partial N} \cdot \tfrac{N}{\tau} = \tfrac{\partial \tau}{\partial S} \cdot \tfrac{S}{\tau} \quad \Rightarrow \quad \varepsilon_N = \varepsilon_S. \tag{C.5}$$

Intuitively, the most efficient charging network for a given level of spending must satisfy the condition that the elasticity of time savings from adding additional stations ε_N should be equal to the elasticity of time saving from increasing the charging speed of the network ε_S.

Notably, the solution does not depend on κ, so we can trace out the efficient network configuration without any further assumptions on the fixed cost structure. To solve for the efficient charging network configuration, we only need to obtain estimates ε_N and ε_S.

C.5 Estimation of the Refueling Time Function $\tau(N, S)$

We use our EV refueling model discussed in Appendices C.1, C.2, and C.3 to evaluate the excess refueling times—the $\tau(N, S)$ function—at a grid of different values of different values of S and N. Specifically, we solve for the excess refueling time across each year between 2011-2021 which provides substantial variation in the number of stations because many new EV charging stations were entering the market over this time period. We also solve the model separately across different assumed charger speeds ranging from 20 kW to 260 kW. In total, we solve the model across 99 (11×9) different possible N, S network combinations. The contour surface of $\tau(N, S)$ based on these 99 points is depicted in Figure 4.

Next, we approximate the surface of $\tau(N, S)$ using the following flexible trans-log functional form fitted to the set of grid points for which we evaluated EV refueling times.

$$\log(\text{Excess Refuel Time}) = \beta_0 + \beta_1 \log(\text{N}) + \beta_2 \log(\text{S}) + \beta_3 \log^2(\text{N}) \tag{C.6}$$

$$+\beta_4 \log^2(\text{S}) + \beta_5 \log(\text{N}) \times \log(\text{S}) + \varepsilon \tag{C.7}$$

The regression results from our main specification are shown in the second column of Figure C.1.[36] We can see that the regression functions provides an excellent fit with an R^2 of 0.987. The coefficient estimates indicate that refueling times are decreasing at a decreasing rate in both N and S.

We use the approximation of $\tau(N, S)$ to evaluate the time elasticities with respect to N and S for the observed EV charger network. Namely, the elasticities are calculated by differentiating Equation (C.6):

[36] The first column shows the results using a simpler "Cobb-Douglass" functional form.

$$
\begin{aligned}
\varepsilon_N =& \beta_1 + 2\beta_3 + \beta_5 \log(S) \\
\varepsilon_S =& \beta_2 + 2\beta_4 + \beta_5 \log(N)
\end{aligned} \tag{C.8}
$$

We see in Table 5 that in 2021, the elasticity of charging speed (S) is nearly 5 times larger than the elasticity with respect the stations (N).

Table C.1: Translog Regression Fit of Excess Refuel Time Surface

	Log(Excess Refuel Time)	
	(1)	(2)
Intercept	6.172***	9.863***
	(0.0793)	(0.2202)
Log(N)	-0.2176***	-0.7133***
	(0.0095)	(0.0483)
Log(S)	-0.3234***	-1.431***
	(0.0120)	(0.0654)
Log^2(S)		0.1172***
		(0.0067)
Log^2(N)		0.0371***
		(0.0036)
Log(N) $\times$ Log(S)		0.0163***
		(0.0053)
Observations	99	99
R^2	0.92850	0.98708
Adjusted R^2	0.92701	0.98639

Notes: Table reports regression estimates of excess refueling time (per stop) on the number of stations (N) and the charger speed of the network. We use 99 different combinations of stations (N) and kW charger speed (S) to fit the regressions. Our preferred specification, Column (2), is used to evaluate the elasticity of excess refueling time with respect to changes in the number of stations (N) and changes to the charging speed (S) of the network.

C.6 EV Refueling Results - Sensitivity Analysis

We make a number of key behavioral assumptions (see Section C.3) to calculate the EV refueling times across different charging network configurations. In this section, we investigate how changes to these underlying assumptions affect the final results. We perform two distinct sets of sensitivity analyses. In the first set, we vary drivers' assumed preferences over the way that they would travel to refuel their EV. In the second set of sensitivity checks, we vary the assumed frequency (and purchase quantity) that drivers would refuel their EV.

Table C.2: EV Refueling Results - Sensitivity Analysis

	Excess Time Per Refuel (Normalized)						
	Total (Min.)	Drive (Min.)	Walk (Min.)	Wait (Min.)	Walk Share [0, 1]	Excess Time Elasticity w.r.t. # of Stations	Excess Time Elasticity w.r.t. Charger Speed
Baseline							
Walk or Wait to Minimize Time	30.63	-0.12	30.10	0.64	1.00	-0.11	-0.52
Refuelling Preference							
Always Walk to Destination	30.63	-0.12	30.10	0.64	1.00	-0.13	-0.47
Always Wait at Charger	532.90	3.91	0.00	528.99	0.00	0.00	-1.00
Refuelling Frequency							
2X Fewer Refuel Stops than Gas	17.63	-0.00	17.29	0.34	1.00	-0.10	-0.95
2X More Refuel Stops than Gas	63.56	-0.02	61.52	2.06	1.00	-0.24	-0.28
10X More Refuel Stops than Gas	414.24	25.42	152.48	236.34	0.58	-0.09	-0.66
20X More Refuel Stops than Gas	618.32	113.06	42.12	463.14	0.13	-0.09	-0.79

Notes: The table shows the sensitivity of refueling choices and the estimated excess refueling times to changes in behavioral assumptions. In our baseline simulation (Row 1) drivers are assumed to refuel EVs at the same frequency (i.e. number of stops per week) as they refuel the gas vehicle and we assume that upon refueling, drivers choose to either wait at the charging station or walk to their destination to minimize total excess time. In the lower rows we show how the results change if drivers prefer to "wait" as opposed "walk" and if driver were to change the frequency they refuel an EV relative to a gas vehicle. The excess time columns indicate the average amount of total excess time to refuel. All times are normalized to measure the excess time per the energy equivalent of a "gas" refueling stop. The walk share column indicates that the fraction of refueling stops that drivers park at the charging station and then walk to their final destination. The last two columns show he elasticity of excess time with respect to changing the number of stations and charger speed, based of 2021 EV charging network.

The first row of Table C.2 shows the results corresponding to our "baseline" assumptions. Namely, we assume that drivers choose the refueling option between "walk" or "wait" to minimize excess time refueling. Moreover, in the baseline counterfactuals, we impose that drivers would refuel their EV at the same frequency as they refuel a gas car, and with energy equivalent "fuel" quantities.

In the baseline counterfactual, we estimate that drivers would spend approximately 31 excess minutes for each time that they refuel their EV. The charging network in 2021 features relatively slow charging speeds (29 kW) so drivers find that walking from the charging station to the destination is time-minimizing nearly 100% of the time. In the baseline counterfactual, almost all of the excess refueling time comes from time that drivers spend walking to and from the charging station. Drivers save 0.12 minutes of driving time each time that they refuel because the chosen charging stations are slightly closer to the trip origin, on average, and drivers also spend an additional 0.64 minutes waiting for the charge cycle to complete after they walk back to their vehicle.

In the middle section of Table C.2 we investigate how the results would change if we assume drivers prefer to always "walk" (row 2) or to always "wait" (row 3). The case where drivers always "walk" is nearly identical to the baseline results because drivers almost always find it to be time-minimizing to walk. In row 3, we see that always waiting at the charger would substantially increase excess refueling time to 533 minutes per refueling stop. This result is explained by the slow charging speed of the current network—at 29 kW charge speed, it takes over 500 minutes to refuel an EV with 256 KWh of electricity (the equivalent of 7.65 gallons of gasoline). In the last two columns of Table C.2, we see that the elasticity of excess time with respect to the number of stations is roughly equal to zero. Although adding more charging stations would slightly reduce driver's time spent *driving* to the charging station, the Table shows that added driving time makes up less than 1% of the total refueling time so therefore additional charging stations would barely change the total excess refueling time estimate. On the other hand, over 99% of the refueling time in this specification comes from waiting at the charging station, therefore any increases in charger speed should reduce the total excess refueling time with an elasticity equal to approximately 1. Hence, if drivers prefer to wait at the charging station instead of walking, this would imply an even larger marginal benefit of increasing the networks' speed compared to increasing the number of stations.

In the bottom section of Table C.2 we solve for counterfactual EV refueling times under different assumptions about drivers' refueling frequency and refueling quantity. More specifically, we calculate excess refueling times for four separate cases. First, we consider the case where drivers refuel their EV half as often as they refuel their gas car while also recharging with double the "fuel" quantity at each stop (i.e., the electricity equivalent of $7.65 \times 2 = 15.3$ gallons). Importantly, increasing the assumed refuel quantity will double the technological time required to recharge the vehicle. Second, we solve the model under the assumption that drivers would stop to recharge their EV twice as often as they stop to refuel their gas cars, and thereby reduce per stop charge quantities by 50% compared to the baseline counterfactual (row 1). Finally, we solve the model for two more cases where drivers refuel the EV 10 times and 20 times more frequently relative to the frequency that they refuel a gas vehicle. To make the excess time estimates comparable across rows, we normalize the excess refueling time estimates and report the excess refueling time per the energy equivalent of an average gasoline refueling stop. For example, if drivers make twice as many refueling stops relative to the baseline, we multiply the refueling time per stop by two when reporting the excess time estimate in Table C.2.

The bottom section of Table C.2 shows that drivers reduce total refueling time from 30.6 minutes to 17.6 minutes when they reduce their refueling frequency by 50%. When reducing refueling frequency, drivers still choose to always to walk from the charging station to the destination, and the vast majority of the total excess time is spent walking. Drivers' are able to reduce excess walking time compared to the baseline case because they are more likely to be able to find charging stations close by to their trip destinations when they refuel less frequently. This result highlights a potential benefit to drivers of increasing the range of EVs. Finally, we also see that the elasticity of excess time with respect to charger speed becomes even larger compared to the baseline. Moreover, the elasticity with respect to charger speed is over nine times larger than the elasticity with respect to the number of stations. Thus, the finding that faster charging speeds would be more valuable to drivers than increasing the number of stations is robust to a case where drivers choose to refuel their EV less frequently that than their gas vehicle.

The last three rows of Table C.2 show that the more often that drivers choose to refuel the EV, the higher the total excess time refueling will be. If drivers' refueling frequency were to double with EVs, then the total excess refueling time roughly doubles to 63.56 minutes compared to the baseline. In this scenario, drivers nearly always find it optimal to "walk" from the charging station to the destination because the technological refueling time is still roughly 3-4 hours. Therefore, this refueling option is strictly dominated by the baseline case (row 1) because total refueling time increases without any reduction in walking. On the other hand, if drivers were to make many more short refueling stops with their EV, then it becomes more likely that drivers would "wait" at the charging station. If drivers increase refueling stops 10-fold then they decide to wait at the station 42% of the time, and if they increase the number of refueling stops 20-fold then they wait 87% of the time. Intuitively, when drivers increase refueling frequency by 20 times and decrease refueling quantity by 95%, the required technological charging time per stop falls below 30 minutes per stop, making waiting more attractive. Although, stopping more frequently makes waiting at the charging stations more practical, we see that overall refueling times are much higher compared to the baseline with less frequent stops but with charging higher quantities.

In summary, we see that the EV excess refueling time estimates are sensitive to the assumptions that we make about driver preferences. However, the comparative result that the excess time elasticity with respect to charger speed is larger than the excess time elasticity with respect to the number of stations hold across all of these alternative assumptions.