2017

A Cost Efficiency Analysis for Private Vehicles: Determinants for Households' Choices of Vehicles Using a Household-Level Commute Data Approach

Tuan Vu
Skidmore College

A Cost Efficiency Analysis for Private Vehicles: Determinants for Households' Choices of Vehicles Using a Household-Level Commute Data Approach

By

Tuan Vu

A Thesis Submitted to

Department of Economics

Skidmore College

In Partial Fulfillment of the Requirement for the B.A Degree

Thesis Advisor: Qi Ge

May 2, 2017

Abstract

In attempts to evaluate the different levels of fuel efficiencies across different types of vehicles, this paper uses a household-level commute dataset to look at the different determinants for a household owning an efficient vehicle. Employing both an OLS and a Probit model, the empirical results illustrate that the current number of vehicles and the vehicle's purchasing price are the attributes that most significantly affect the household's probability to own an efficient vehicle. A similar analysis is adopted for the case of electric vehicles as well. A further analysis includes calculations for different total costs of owning vehicles with different fuel economies. The results of these calculations suggest that while the more efficient vehicle is more expensive to own at first, its benefits will outweigh its costs as the vehicle is utilized more.

1. Introduction

In recent years, given rising concerns about the limited supply of fuel oil, as well as the increasing global effects of carbon dioxide (CO_2) emissions and greenhouse gasses, there has been a trend toward using more fuel-efficient vehicles. According to a 2012 report of the European Environmental Agency, 25% of CO_2 emitted in the European Union came from transportation, with three-fifths of this amount stemming from the use of private vehicles (Kihm and Trommer, 2014). In a world of hectic movement like today, the demand for travel is higher than ever. Over time, we have designed and created newer and faster means of transport, including, but not limited to, trains, subways, and airplanes. With regards to the use of private vehicles, as technology advances, we have new models of vehicles with better utilities as well as better fuel economies. This development is not limited to conventional gasoline vehicles (CGVs), as we have developed newer and more fuel-efficient means of transport for the private vehicle sector, namely the developments of electric vehicles (EVs).

Currently, there are several types of EVs with different mechanics on the market. The most prevalent types of EVs are battery electric vehicles (BEVs), hybrid electric vehicles (HEVs), and plug-in EVs (PHEVs). One common source of energy for all of these types of vehicles is electricity, but the way electricity is generated is different for each type of EV. The details on the mechanisms of these types of EVs will be discussed later in this paper.

Nonetheless, despite the various types of EVs currently available, CGVs are still the type of vehicles that dominates the current private vehicle market. Moreover, as technology advances, the levels of fuel economies for CGVs increase as well. A question arises to every single household when considering buying a new vehicle: How much do I value an efficient vehicle?

Thus, the decision on whether to own an efficient vehicle or not is essentially an economic one: if the benefit of owning an efficient vehicle can outweigh its cost.

In this paper, by using household-level data retrieved from the Panel Study of Income Dynamics (PSID), I formulate a model that calculates the probability of a household owning an efficient vehicle, with a further analysis for owning an EV, based on household characteristics from 2011and 2013. The PSID dataset contains information at a household-level, including the number of vehicles available in the household, the manufacturer/model/type of up to three vehicles in the household, the average daily commute time of both the head and the wife, as defined in the dataset, of the household, the household's monthly gasoline expense, the vehicles' purchasing price, and the household's annual income for the previous year. This dataset offers an inclusive set of variables, because not only that it provides the household's currently available vehicles but it also shows the household's driving habits. Thus, given these types of variables, I can more accurately compute the probability, as well as the cost, of owning an efficient vehicle, as well as an EV, based on a household's travel demands.

The regression outcome in my study provides some noteworthy findings for the determinants for owning an efficient vehicle as well as for owning an EV. Generally, the results for both 2011 and 2013 indicate a higher probability of owning an efficient car as the total daily average commute time increases, with the highest increase of 0.9% for efficient vehicles and 0.2% for EVs, though this effect differs between years and models. Out of all the variables, the number of vehicles available has the strongest correlation with the probability of a household owning an efficient vehicle, to the extent that an addition vehicle can increases the probability of owning an efficient vehicle by 8-9% for 2011 and 2013. However, the same effect is not present in the probability of owning an EV.

Regarding the total cost of ownership (TCO) for different types of vehicles, this study divides the population of vehicles into efficient and non-efficient vehicles, and calculates the TCOs for efficient and non-efficient vehicles separately for both 2011 and 2013. The results indicate that for both 2011 and 2013, overall it is costlier to own an efficient vehicle than it is to own a non-efficient one, despite the fact that these vehicles have higher levels of fuel economies. However, the differences between TCOs for efficient and non-efficient vehicles decrease from 2011 to 2013, with the average differences of approximately $8500 in 2011and $1300 in 2013. Interestingly, there is an overall increase in TCOs for all vehicles from 2011 to 2013, but the increase in TCO for non-efficient vehicles in 2013 is the most noticeable one (from $26190 to $35901 on average). This increase is the result of the increase in travel demands of households from 2011 to 2013. As the non-efficient vehicles have much lower fuel economy, they will incur a much higher operating cost when travel demands increase. A further analysis of TCO is also done for EVs and CGVs using the same method as for efficient versus non-efficient vehicles. However, unlike the TCOs for efficient and non-efficient vehicles, there is a switch in the gap between the TCOs for EVs and CGVs as travel demands increase from 2011 to 2013. Specifically, in 2011, it costs roughly $3000 more on average to own an EV; however, in 2013, it costs almost $8000 less on average to own an EV.

There are two main contributions of this study. The first one is the analyses that determine the probabilities that a household will own an efficient vehicle, or an EV, given different household's attributes. The second main contribution of this paper is the calculation of TCO for vehicles while allowing the households to utilize their bundles of vehicles. Moreover, unlike previous studies that only calculate the TCO of one vehicle, by allowing for vehicle utilization, my study also includes a TCO calculation when all available vehicles within a household are taken into account, thus representing a TCO for all vehicles available, instead of just one vehicle.

The rest of the paper is organized as follows. Section 2 includes a review of existing literature on this topic. This section is divided into two main parts, with the first part primarily focusing on the utilization of vehicles within households and the second part looking at past models used by other researchers to calculate the cost of owning EVs instead of CGVs. Section 3 describes the data retrieved from PSID as well as other supplementary sources. Section 4 discusses the methodology applied in the study. The paper ends with section 5, which interprets and rationalizes the regression results, as well as compares the results with previous studies. The paper ends in section 6, which offers some insights drawn from the results, and concluding remarks where further research ideas are included.

2. Literature Review

When it comes to the topic of efficient versus inefficient vehicles, in most cases consumers make this distinction based on vehicles' levels of miles per gallon (MPG). While it is true that there are CGVs that have a high level of fuel efficiency, EVs are a special case of efficient vehicles, since not only that they can generally have a higher level of fuel efficiency when compared to CGVs, but also that EVs have a lower level of CO_2 emissions due to the reduced uses of gasoline. As the existence of EVs is becoming more omnipresent and well-accepted by consumers, more and more users are considering EVs as a feasible replacement for their existing CGVs. As suggested by Tseng et al. (2013), annual sales of EVs in the U.S. have grown from 1% in 2004 to 4.4% in 2011. Furthermore, in the case of the U.S., due to its massive geographical territory, owning at least one vehicle has become a necessity to almost every single household. EVs were introduced as the more environmentally friendly and fuel-efficient alternative to CGVs. Yet, given the current market condition where the purchasing prices for most EVs still generally lie in the higher price range when compared to that of other CGVs, as well as the range limitations in the present developments of EVs, many users, even the environmentally concerned ones, are deterred from owning an EV either as a primary or secondary vehicle (Hidrue et al., 2011). Along these lines, I find the need to study how households utilize their bundles of vehicles and analyze the prospect of owning an EV as a substitute for a CGV in daily travel commute.

This literature review is divided into two main literature groups. The first group of literature focuses on how households utilize their bundle of vehicles. Following this, I briefly explain the mechanism of currently available types of EVs. The final part of the literature review focuses on describing different methods of calculating the TCO implemented by various studies with regards to EVs. The literature review ends with an overview of where my study stands and how it can contribute to this field of research.

2.1. Household Demands for Vehicles

Before looking at the EV market, it is first important to understand households' vehicle choices and usages. As mentioned, in the context of the U.S., owning at least one vehicle has become crucial to many households, and it is common for a household to own more than one vehicle. Given that households have different travel demands, the ways in which households make their vehicle purchasing decisions are based on their travel demands. Intuitively, as there are multiple members in the households who have needs to travel using vehicles, the higher the total households' travel demands become, and the more vehicles are purchased. Furthermore, as families increase the number of vehicles available in their households, the ways in which households can utilize their bundles of vehicles increase as well. Since households' travel demands can be reflected in their choices of vehicles, it is important to understand the factors that influence households to make their vehicle purchasing decisions, as well as how households utilize their choices of vehicles.

Many studies have tried to explore the factors that contribute to how households make vehicle purchasing decisions. In their research, Bento et al. (2005) look at the effect that different urban forms have on how households choose their vehicles and how these vehicles are utilized. Using the data from the 1990 Nationwide Personal Transportation Survey (NPTS), they construct a dataset that includes approximately 20,000 U.S. households in 114 different urban areas. The data offers information on household's characteristics, such as income, race, gender, education, etc., as well as the household's choice of vehicles and the annual miles driven. They then construct two models: one looking at how the mode of commuting is chosen, and the other focusing on explaining the number of current vehicles and the miles driven per vehicle.

In the commute mode choice model, Bento et al. (2005) look at how household characteristics can affect the household's choice of commutes. The different modes of commute considered in the study are driving, walking/biking, taking the bus, or taking the train. Using the NPTS sample, they find that, similarly to previous literature, income, race, and education all significantly affect the commuter's choice of commutes. Not surprisingly, it is found that workers with higher income are less likely to take public transportation or walk to work, as they are more likely to be able to afford a car. Race also plays an important factor in the sense that white people are the ones least likely to take public transport.

Household's characteristics aside, population centrality, defined as the percentage of households living near the center of the area, also has a significant effect on the probability of the commuters choosing whether to drive or to take public transportation. According to their regression results, Bento et al. discover that a 10% increase in population centrality can lower the probability of choosing to drive to work by 1%. This can be translated to a reduction of 54 miles annually assuming the average annual miles driven of a worker is 6000 miles (Bento et al, 2005).

In the second model, this study focuses on the determinants of the number of current vehicles in the household and the household's demand for vehicle miles traveled (VMT) per vehicle. Using the same sample, Bento et al. (2005) find that household size has a significant effect on the probability that the household will have an additional vehicle. According to the model, on average an additional working member to the household can increase the household's annual VMT by approximately 5000 miles, with 4000 of which are the result of the additional number of vehicles. Thus, it can be concluded from this result that an addition of a working adult has a much greater effect on the increase of the number of vehicles in the household, significantly more than the effect it has on the annual VMT per vehicle.

In their study, Bento et al. (2015) focus on the factors affecting households' commuting choices. My study will advance one step further by using household's commuting time as a variable that represents the household's travel demands to explain their choice of vehicles, with an addition of EVs. Instead of looking at all modes of transportation, my study will narrow down to only private vehicles as the primary mode of transportation for households. Furthermore, similar to how Bento et al. (2015) focus on the determinants of owning an additional vehicle, my study looks at the determinants of owning an efficient vehicle, as well as an EV, by forecasting the probability of a household owning an EV based on factors similar to Bento et al.'s (2015), such as the household's income, the number of existing vehicles, and household's commuting time. Unlike Bento et al. (2015), since my study is taking the aggregate commuting time of the household, household size will not be employed in my study. However, based on Bento et al.'s (2015) result that there is a positive correlation between the number of working adults and household's annual VMT, I will use annual VMT based on commuting time as one of the determinants to forecast the probability of owning an EV, as well as to calculate TCO for all vehicles in the household.

Many other studies choose to employ a discrete-continuous model to study household's choices when utilizing their vehicles (Spiller 2012; Fang 2008). In the classical discrete utility choice model, which is used to model the utility one gets based on that person's decision, there is an assumption that the choices are made independently of one another. When the choices are assumed to be made independently from one another, there lies a further assumption that there is no diminishing marginal utility associated with the current choice when the level of consumption of any other choices increases. Yet, this is not the case for owning a bundle of vehicles. When the number of vehicles in the household increases, there is diminishing marginal utility in choosing to utilize a vehicle since driving one vehicle would result in an opportunity cost of not driving the other vehicle (Bhat, 2005). Thus, in order to account for this diminishing marginal utility, Bhat (2005) derives a model based on the classical utility theory for discrete and continuous choices. For households that own more than one vehicle, the choice of which vehicle to drive occurs simultaneously between multiple alternative vehicles. Unlike the classical discrete utility function where only one alternative is chosen from a set of mutually exclusive alternatives, the multiple discrete-continuous function deals with situations where consumers deal with multiple alternatives, which in the cases of transportation research are the other available vehicles, simultaneously. The model is derived from adapting a translated non-linear form of the utility function from previous research, with an addition of a multiplicative log-extreme value error term (Bhat, 2005). However, in his study, Bhat (2005) does not implement the multiple discrete-continuous model to explain the individual's vehicle utilization decisions, but to explain how the individual spends time in different types of activity pursuits. Individual activity pursuit is very similar to vehicle utilization decisions, in the sense that both situations involve a choice being picked from a range of multiple alternatives occurring simultaneously. As suggested by Bhat (2005), the multiple discrete-continuous model can also be applied in the context of vehicle utilization, as done by other studies such as Spiller (2012) and Fang (2008).

As mentioned, in transportation studies, the application of Bhat's (2005) multiple discreet-continuous function is widely used. In most real-world situations, decisions are not made independently from one another, but instead some decisions are interconnected and required to be taken simultaneously (Ahmand et al, 2015). By implementing the discreet-continuous model, this interconnectedness of decisions can be accounted for. When it comes to the decision of purchasing or utilizing a vehicle, there are multiple factors involved in this decision-making process, such as how the vehicle will be utilized by the household, based on the household's commuting habits, and how much the cost of fuel will be given the household's travelling habits. Thus, researchers find a need to implement the discreet-continuous model when it comes to transportation research. In the application of the discreet-continuous model to transportation studies, the type of vehicles chosen by a household is the discreet variable, as there can only be a finite numbers of vehicles available, and how the vehicle is utilized, in other words, how many miles each vehicle is driven by the household, is the continuous variable (Ahmad et al., 2015).

In order to determine the effect that residential density has on vehicle choice, Fang (2008) employs the multiple discreet-continuous model derived by Bhat (2005), but to analyze it in the context of vehicle utilization using a dataset that includes vehicle properties (such as price, mpg, etc.). Using data from the National Household Travel Survey (NHTS) in 2001, Fang (2008) finds a negative relationship between density and the number of cars or trucks in the household. Somewhat similar to Bento et al.'s (2005) results, Fang (2008) finds that as the area the household lives in becomes denser and more centralized, the probability of driving to work decreases, and thus in the long-run, households will eventually reduce the current numbers of vehicles available in their garages. In addition to the multiple discreet-continuous model, Fang (2008) also proposes a method using Probit and Tobit models to analyze household decisions on the number of vehicles. Since Fang (2008) is interested in looking at the probability of owning a certain number of vehicles with respect to changes in population density. Probit and Tobit models are implemented since they both derive the likelihood of the dependent variable occurring based on the given independent variables. A small difference exists between these two models is that while the Probit model can show the signs as well as the probabilities for the independent variables with regards to the dependent variable, the Tobit model is designed to estimate the actual change in the dependent variable above a certain threshold. In general, the two models are very similar mathematically, with the Probit model being less sensitive to the distributions of specifications (McDonald and Moffitt, 1980).

Fang's (2008) study implements these two methods to specifically look at the probability of the household holding certain numbers of vehicles as density increases. In the case where density increases by 50%, she finds that the change in probability for a household choosing a truck is negative, while this change is positive for choosing a car. Thus, it can be said that households view trucks and cars as substitutes as density increases (Fang, 2008). When the two methods are compared, a consistency is found in both with regards to miles travelled as density increases.

Similar to Fang (2008), using the same discrete-continuous model derived from Bhat (2005), Spiller (2012) employs a utility function that uses the same NHTS data from 2001 that Fang (2008) uses, with an addition of the year 2009, to specifically looks at gasoline demand with respect to vehicle utilization. Spiller (2012) argues in her research that past studies have not accounted for households' bundles of vehicles when calculating elasticity of demand for gasoline, so in her model Spiller (2012) accounts for how much people drive (based on VMT) and what types of vehicles they have. She finds that the elasticity of demand for gasoline is -0.89, which indicates that the demand for gasoline is inelastic as it is less than 1. However, when compared to the case where elasticity for gasoline demand is computed independently among vehicle choices, the elasticity of demand for gasoline in the discreet-continuous case is higher (-0.89 compared to -0.62). Thus, this result confirms that allowing households to optimize their choice from their vehicle bundle increases the elasticity of demand for gasoline. It is also suggested in this paper that by not allowing for the utilization between vehicles, past research has underestimated the elasticity of demand for gasoline by up to 66% (Spiller, 2012).

In my study, in addition to using commuting time, I will also be using gasoline expenditure as one of the determinants that affects the probability of owning an efficient vehicle as well as an EV. Moving forward from Spiller (2012) who looks at the elasticity of demand for gasoline as VMT changes, my study will take into account both the changes in gasoline prices and in VMT to predict the probabilities of a household of owning an efficient vehicle and an EV, as well as to calculate the TCO for vehicles, while allowing for utilization between different types of vehicles.

In her analysis, Spiller (2012) also makes an observation that in reaction to a change in gasoline prices, in the short-run households can drive each of their vehicles less, and eventually reallocate their driving patterns to optimally utilize their bundles of vehicles. Acknowledging Spiller's (2012) attempt to allow for substitution between vehicles within the same household, Borger et al. (2014) focus on how the change in gasoline prices influences multi-vehicle households' driving habits, especially looking at how households substitute their choices of vehicles with regards to fuel efficiency. Assuming that the primary vehicle in the household is the one being used the most during the period of observation, Borger et al. (2014) theorizes that there is a substitution effect towards the most fuel efficient car available in the household. Since this study is strictly looking at the substitution effect of vehicles within multi-vehicle households, the samples considered in this study are only the households with two vehicles. The results are in accordance with their hypothesis, that for a given change in gasoline prices, the less fuel-efficient car will incur a higher change of cost per kilometers. Thus, given this increase in gasoline price, households will eventually shift their driving toward the more fuel efficient car (De Borger et al., 2015).

The literature on vehicle utilization is relevant to my study as my study focuses on how households make vehicles purchasing decisions and the probabilities of a household owning an efficient vehicle and an EV. When looking at whether an efficient can be a substitute or a complement to a non-efficient vehicle, it is first important to understand the factors that households base their decisions on when making a vehicle purchasing decision as well as their vehicle utilization decisions. The literature mentioned above explains the influences different factors have on these decision-making processes. Moving forward from Fang's (2008) study, my study specifically looks at the probabilities of a household owning an efficient vehicle, as well as

EVs. With an addition of EVs, my study can further add more choices of the bundles of vehicles given to consumers. By allowing EVs as another vehicle option, given the differences in prices as well as levels of fuel economy, households will have more room to utilize their bundles of vehicles and minimize their costs of transportation.

2.2. Different Types of Electric Vehicles

First, it is important to note the differences in various types of EVs. EVs generally consist of battery electric vehicles (BEVs) and hybrid electric vehicles (HEVs). BEVs are the type of EVs that use the energy generated from a stored rechargeable battery. Some commonly known models of BEVs that are currently available on the market are the Nissan Leaf, the Tesla Model S, etc. Unlike the purely electric BEVs, HEVs are the type of vehicles that combine the energy generated from a conventional gasoline combustion engine and an electrical generator. For HEVs, the electrical energy is generated from a technology called regenerative brake, which allows the HEV to make use of the kinetic energy from braking and converts this energy into electric energy. The Toyota Prius is the most commonly known HEV on the current market. There is also a special type of HEV, called plug-in hybrid electric vehicles (PHEVs), which are essentially EVs that use the combined energy from electricity and gasoline, but instead of using the special braking system to generate energy, the PHEV, similar to the BEV, has a rechargeable battery inside the car, which allows the car to be recharged when plugged into an electrical source (Alternative Fuels Data Center, 2016). Some well-known PHEV models are the Chevrolet Volt or the BMW i8. Besides those, many other HEV models, such as the Toyota Prius, also have a PHEV option. Given these differences in their operating engines, different types of EVs also have different levels of economy.

2.3. Past Total Cost of Ownership Models

In an attempt to accurately measure the true TCOs for both CGVs and EVs, numerous studies have tried to implement a variety of models to precisely capture and compare the entire scope of TCOs for different vehicles. A basic TCO calculation for vehicles includes the fixed cost, which is mainly the purchasing costs of vehicles, and the variables costs, which include the operating costs as well as other external costs, such as maintenance costs.

In their paper, Rusich and Danielis (2015) construct their own TCO model based on a case study conducted in Italy. Using the data collected from 66 different car models in their case study, they formulate a TCO model that can account for the capital cost of the vehicles as well as annual operating costs. Out of the 66 vehicles surveyed, there are 10 models of vehicles for CGVs as well as for HEVs, and 14 models for BEVs, with 4 of those being BEVs with leased battery; the rest of the vehicles are either vehicles that use diesel or natural gas. This paper outlines the basis of a TCO calculation, which includes costs such as vehicle capital cost, annual capital cost, average annual insurance cost, annual maintenance and repair cost. For EVs, they also include an annual electricity cost, which is deconstructed into different levels of fuel efficiency based on primary and secondary fuel ranges for the cases of HEVs, as well as an annual battery leasing fee for EVs that require leased batteries. Rusich and Danielis (2015) then calculate the TCOs for different vehicles given different variations in the annual kilometers driven, ranging from 15,000km to 25,000 km per year. Their results show that in Italy after a 5-year interval, CGVs have the lowest TCO when compared with HEVs, PHEVs, and BEVs. It is also noted by Rusich and Danielis (2015) that BEVs can become convenient only when the annual distance traveled is at least 20,000 km. (Rusich and Danielis, 2015). The one flaw in Rusich and Danielis's (2015) study is the assumption that all vehicles have the same total annual distance driven. This assumption is sensible when only the vehicles' attributes are looked at to calculate TCOs, but when looking at the determinants of owning an EV, it would be inadequate and inaccurate if the drivers' driving habits were excluded.

Similarly, with the combination of data retrieved from the Bloomberg New Energy Finance and Eurostat in 2010, Thiel et al. (2010) construct a similar TCO model for EVs that accounts for the vehicles' purchasing costs as well as several powertrain costs and battery costs, for EVs, under the assumption that the average annual mileage of a passenger car is 15,000 km (according to an approximation by the European Automobile Manufacturers' Association in 2008). When compared to the previous study done by Rusich and Danielis (2015), not only the inclusion of PHEVs is introduced in Thiel et al.'s (2010) paper, but also the inclusion of a payback period analysis, where the learning effects of technology is accounted for. Noticing the currently fast growing technology in the EVs market, Thiel et al. (2010) assume a faster learning rate for EVs, at 10%, than for CGVs, at 5%. At the current market level, they find that when compared to CGVs and diesel vehicles in 2010, which have a payback period of 6-7 years, BEVs can have a payback period of about 22 years on average, with similar periods for both PHEVs (22) and HEVs (20) in 2010. The reason for this gap in payback period is due to the difference in high purchasing costs, which will still remain an issue until federal support is provided. However, after applying the rates of different learning effects mentioned above, there is a huge decrease in payback period for all classes of EVs. Specifically, in 2020, it is forecasted that while the payback period for CGVs does not change much, BEVs and HEVs will likely to have a payback period of more than 8 years, with PHEVs being the ones with the longest payback period out of all EVs' classes with 10 years. (Thiel et al. 2010). Thiel et al. (2010) explain that this reduction in payback periods is the result of technology cost reductions, which can be achieved through learning effects. However, the same issue as Rusich and Danielis's (2015) study is present in Thiel et al.'s (2010), which is the assumption that the drivers' annual mileage does not change, even after a long period of time. While these studies have successfully identified the relevant variables in the calculations of TCOs as well as payback periods for vehicles, they have not accounted for drivers' actual driving habits, which can be very influential when it comes to the determinants of purchasing and utilizing a vehicle from the consumer's standpoint.

However, Wu et al. (2015) argue that the studies regarding TCO for EVs in the past, such as those similar to Thiel et al.'s (2010) and Rusich and Danielis's (2015), only mention the aspects regarding the vehicle but do not account for the drivers' driving habits. Many researchers explain that due to the variance in our daily travel demands, it is necessary that the driver's driving habits should be considered in order to more accurately assess EV's efficiency and its TCO. In order to solve this problem, studies have tried to analyze the TCOs for vehicles using a GPS travel data approach, which provides information on the actual daily driving patterns. By doing so, they can more accurately capture the TCO for different vehicles given various driving habits (Wu et al., 2015; Wu, Aviquzzaman, & Lin, 2015; Li et al., 2016). Given their travel data, which is taken from a report filed by the German Federal Motor Transport Authority in 2012, Wu et al. (2015) calculate and compare the TCO per km for CGVs, HEVs, PHEVs, and BEVs. Similar to past studies, Wu et al.'s (2015) TCO analysis consists of the initial purchase cost of the vehicles, its resale values, and annual operating cost with respect to discount rate. This study advances one step further by looking at the TCO divided by the annual kilometers traveled. By breaking the TCO of the vehicles down to a kilometer scheme, Wu et al. (2015) illustrate the TCO for a broad range of different types of vehicles with a more accurate and detailed view at

TCO. When looking at the starting year, which is 2014, CGVs are the vehicles with the lowest TCO, at about 42 cents/km for the medium range. However, in their comparison of the mean TCO/km over a 10-year period, Wu et al. (2015) find that as time increases, HEVs eventually have a lower TCO than CGVs, while PHEVs and BEVs still have much higher TCOs. Furthermore, it is suggested by the results that the difference in cost between a CGV and a BEV decreases from 12 cents/km to 3 cents/km. They also conduct a forecast in which they find that, in 2025, HEVs will have a probability of 51% to become the vehicles with the lowest TCO/km; that number is 40% for CGVs.

This study also divides vehicles into different classes with respect to their size and analyzes these vehicle classes separately at different driving parameters. In general, the result of this study indicates that across all classes of vehicles as well as different driving ranges, EVs have a higher level of cost efficiency relatively to CGVs. However, their model indicates that in the short distance range, CGVs are more likely to be the more cost efficient vehicle, but the opposite is true for both the medium and long range distances. Nevertheless, the results do not provide a strong enough significance to clearly distinguish which one would be the most cost efficient one.

Another study similar to Wu et al. (2015) is conducted by Bubeck et al. in 2016. This study implements the same method and the same dataset as Wu et al (2015) do. What distinguishes Bubeck et al.'s (2016) study from Wu et al.'s (2015) is that, instead of separating vehicles into different size segments, Bubeck et al. (2016) focus more on consumers by dividing the sample into different user types with different annual mileages. Unlike Wu et al. (2015) who use actual driving data, Bubeck et al. (2016) divide the users into three types: low mileage driver (7500 miles annually), medium mileage driver (15000 miles annually), and professional high mileage driver (75000 miles annually). Their result is also different from that of Wu et al. (2015), as Bubeck et al. (2016) find that professional high mileage drivers can have the lowest TCO if using a CGV. Nevertheless, simply grouping users into different mileage group can only reflect a broad estimation of TCO and does not necessary illustrate the true TCO compared to when actual driving data is used. While Bubeck et al.'s (2016) study has advanced one step further from the studies done by Rusich and Danielis (2015) and Thiel et al. (2010) by dividing drivers into different groups based on annual mileage, yet, the same issue is encountered, as the assumption that the driven mileage of drivers is constant exists in their study. Thus, Wu et al. (2015) is the only study that most accurately calculate for TCOs of different vehicles while accounting for drivers' driving habits.

Further literature also argues that even though past models have focused on projecting a long-term TCO for owning a vehicle, yet they do not take into account the fact that many people have the tendency to change their vehicles after several years of usage. It has been noted by Gilmore and Lave (2013) in their study that many owners sell their vehicles after three to five years of usage. Noticing this trend in the current used car market in the U.S., Gilmore and Lave (2013) construct their TCO model by using the resale prices of vehicles. They assume that when asked to choose between two vehicles of equal attributes in all aspects except for the type of fuel, the rational consumer would only buy the vehicle with alternative fuels (HEV, PHEV, or BEV) if the vehicle's fuel saving cost can recover its initial high purchasing cost. After grouping the vehicles with the closest attributes into pairs of twos, the study compares the difference in resale prices and expected fuel costs. The result indicates that for the pairs of passenger vehicles, the diesel and HEV options have a lower TCO when compared to CGVs, but in the pairs of larger vehicles (SUVs, for instance), the diesel option has the lowest TCO. However, a limitation of this study is that the scope of vehicles studied is limited only to those available in the resale markets. Since the EV market is still currently in its developing phase, as suggested by Thiel et al. (2010), EVs will likely to have a much faster learning effect. Thus, there will be many new EV models with better fuel-efficient technologies available for sale in the primary market, but these vehicles will not be available in the resale market just yet. Hence, this TCO model fails to capture all currently existing models, making it inadequate to accurately evaluate all of the current models of EVs.

When looking at the prospect of EVs in multi-vehicle households, Tamor and Milačić (2015) estimate the acceptability of EVs by analyzing actual one-day travel distances of households in the Seattle area, using the data retrieved from the Puget Sound Regional Council Traffic Choices Study in 2008, which was made available by the National Renewable Energy Laboratory in 2013. According to their estimation, Tamor and Milačić (2015) find that the notion of EVs would be much more accepted if the household owned two vehicles. For single-vehicle households, a replacement of an EV simply does not solve the fuel-efficiency issue due to the limitations of driving range. For multi-vehicle households, this is no longer the case. By only substituting one of the vehicles with an EV option, the households can optimize their bundle of vehicles as the level of inconveniences decreases with a higher overall level of fuel-efficiency. As driving range increases for EVs, so does the level of fuel-efficiency, while the inconvenience stemmed from the range issue decrease.

As mentioned, in the U.S., the notion of owning more than one vehicle is no longer foreign to households, and thus simply calculating the TCO for one vehicle at a time cannot fully capture how users utilize their vehicles, as well as the TCOs for households that own multiple vehicles. The primary purpose of this paper is to study how households utilize their vehicles with respect to their daily travel commute habits, with an addition of EVs. Moreover, the TCO calculations in this study are not limited to the calculations of only one vehicle, but instead the TCO calculations here can capture the entire cost of owning multiple vehicles, while accounting for how households can utilize their bundles of vehicles. Thus, it is necessary to consider both how households utilize their vehicle usages and TCOs for vehicles, with a specific case for EVs. Past studies have looked at the issue of vehicle utilization within a household and the different levels of fuel efficiency for EVs separately, but only a few have incorporated the vehicle utilization decisions when calculating for the TCO of EVs. One of those few studies is the study conducted by Tamor and Milacic (2015), but it only looks at the prospects of EVs in multi-vehicle households, and yet does not take into account how households can simultaneously utilize EVs and CGVs differently. Thus, this paper intends to bridge the gap left by prior studies, which is the utilization of vehicles within a household with an addition of EVs.

Furthermore, most of the TCO studies presented above were done in Europe, while a very few has been conducted for the case of the U.S. Being a country where public transportations are not as accessible at a nationwide level as most countries in Europe, the U.S. is heavily dependent on the uses of private vehicles when it comes to transportation. By using the same method applied by Wu et al. (2015), I will calculate the TCOs for different classes of vehicles with respects to VMT based on workers' commuting time in different states of the U.S., while account for vehicle utilization decisions. Unlike previous studies which only calculate the TCOs of one vehicle at a time, by allowing for vehicle utilizations within a household, my model calculates the TCOs for the households owning all of their available vehicles, with respects to their commuting habits and vehicles' purchasing prices. Additionally, I will look at the determinants in which households base on when making a fuel-efficient vehicle purchasing decision by analyzing the probability of owning an efficient vehicle given different schemes as suggested by Bento et al. (2005), Wu et al. (2015), and Rusich and Danielis (2015), such as the total number of vehicles in the household, gasoline expenditures, total commuting time, the household income of the previous year, vehicles insurance payments, as well as purchasing prices of vehicles. In order to allow for more vehicle options based on their fuel types, a similar analysis is done for the case of EVs. As suggested by Spiller (2012), consumers are more sensitive to costs in order to minimize costs with the inclusion of vehicle utilization, I hypothesize in my study that by allowing for vehicle utilization decisions, the TCOs for vehicles will be less than when vehicle utilization is not allowed.

In short, this paper aims to study the determinants based on which households make their vehicle purchasing decisions, and calculate the costs of households owning different vehicles given different levels of vehicles' fuel efficiency and households' travel demands. In addition to calculating the TCO for only one vehicle at a time, by allowing for vehicle utilization decisions, this study can calculate the entire TCO for all vehicles available in a household.

3. Data description

The main source of data used in this study is taken from the Panel Study of Income Dynamics (PSID) from the University of Michigan, with other supplementary datasets taken from the U.S. Department of Energy, the Governors Highway Safety Association (GHSA), and the Energy Information Administration (EIA) dataset. The PSID offers a household-level dataset for households living in the U.S., and this study focuses primarily on variables that are related to the attributes of the vehicles in the households. Due to the availability of data, the period of interest for this study are the years 2011 and 2013.

3.1. PSID

At a family level, the PSID provides details on the current vehicles in the households such as the number of vehicles available in the households, with details on the manufacturer, make, brand, year of the car and the price it was purchased at, as well as a hybrid indicator showing whether that specific vehicle is an EV or not. For these variables, PSID provides specifications for up to three vehicles in the households, labeled as vehicle one, vehicle two, and vehicle three respectively. By having the actual purchasing price of the car, whether the car is a used or a new car will be accounted for. This dataset also includes these same variables for the other two vehicles in the household, when applicable. Other control variables taken from the PSID include the average daily commute time, in minutes, of both the head and the wife of the family, the total expense of gasoline for the previous month, the amount of insurance paid per corresponding periodical interval, and the family total income for the previous year.

Since I am interested in the total annual cost of owning a vehicle, all periodically controlled variables are converted into an annual term. The average daily commute time variable is defined as the total minutes it takes the household's head or wife to travel a round trip commute to and from work on a typical day. First, since the commuting time for the head and the wife are separated, but the vehicle used to commute by each was not specified in the dataset, I combine them together to get the total average commute time for the household. By combining the average commute time of the head and wife, I can also identify and account for those households where the average commute time is zero for both the head and the wife, as well as the ones where average commute time is only zero for either the head or the wife. Thereupon, all the observations where the total average time of the household is equal to zero, meaning those for which both the head and the wife of the household have zero commuting time, are dropped. I then divide this variable by 60 in order to get the average daily roundtrip commute time in hour terms, followed by multiplying this number by 261 days. The number 261 stems from the fact that there are 52 weeks in a year for a total of approximately 104 weekend days, and since this study is mainly concerned with commuting time to and from work, only weekdays will be taken into account. There lies an assumption here that no vacation days or national holidays were used by the household. Thus, this computation allows me to generate a variable that indicates the annual average commute time of the household in hours.

Another controlled variable that needs to be converted into an annual term is the insurance expense variable[1]. The PSID offers two separate variables for insurance expense: one is the actual amount of insurance paid by the household for all vehicles in monetary terms, and the other is a time unit variable that displays the period per which the insurance was paid in, either monthly or annually. Accordingly, I generate an annual insurance expense variable, which is computed as the product of the monetary insurance expense variable and the time unit periodic insurance payment variable, with 12 for monthly payments and 1 for annual payments. Furthermore, since the insurance payment is the total amount for all available vehicles, I calculate the average amount of insurance paid for each vehicle by dividing the total amount of insurance over the number of vehicles available in the household. All observations where the amount of insurance expense is not specified are dropped. The same method is applied to the total amount spent on gasoline for transportation related expenses. Since this variable is already

[1] Note that the insurance expense was not included at first, but it was merged with the master dataset afterward, and this generated 15 missing values where the households did not indicate an insurance expense

in monthly terms, I multiply this expense by 12 to get an annual expense on gasoline for transportation. It is assumed here that the total expense on gasoline is counted as the cost of traveling to and from work, so any gasoline expense on transportation for leisure or vacation is not included. The family total income is already an annual term, so this variable does not need any further changes.

When looking at the vehicle attributes within households, I first identify and drop all the observations where the number of vehicles in the household are zero or unidentified. That way, the pool of observations is limited to only households that own at least one vehicle. Thereafter, because a single vehicle is identified by multiple variables, namely manufacturer, brand, year, and hybrid indicator, I create a single variable acting as a vehicle identifier which can capture all of these aforementioned attributes. In the PSID, these identifying variables are denoted as numbers, with a general variable for the manufacturers (e.g.: 32 for Toyota), and another more specific variable used for the brand of the car (01 for Toyota and 02 for Lexus if the general variable was 32). The identifier is generated by compiling all the attributes of the vehicle into a single series of numbers that can distinguish that car. For instance, using the same Toyota example, a Toyota car is identified as 321, while a Lexus car is identified as 322. I further compile this variable with the model's year by simply putting the year following the above series of manufacturer and brand (so a 2010 Toyota will be identified as 3212010). For the sake of simplicity as well as the availability of data, any model of vehicle prior to 2000 is classified in the same group as those in the year 2000. Lastly, the hybrid indicator is a dummy variable, with 1 indicating that a car is hybrid and 0 not a hybrid. The PSID has two indicators for this dummy. They are both questions that ask whether a certain vehicle is a hybrid or not, although the first one is asked in the case that the model is known; whereas the other is asked only when the model of the car is unknown. Before compiling this indicator into the vehicle identifying variable, I first generate a new dummy hybrid indicator that accounts for both the vehicles with known models and the unknown ones by setting the new dummy variable equal to 1 as long as either one of the original PSID indicators is equal to 1. Applying the same method, I put the newly generated hybrid indicator at the end of the number string that identifies a vehicle's manufacturer, brand, and model year. To be consistent with the example above, a 2010 hybrid Toyota is identified as 32120101, while a non-hybrid one is identified as 32120100. Notice that for the year 2013, there is an addition of BEVs in the original PSID hybrid indicator, denoted as 2. To simplify this, I group the vehicles that are electric and hybrid into one category (henceforth referred to as EV), defined as 1 in the newly generated hybrid indicator. This method was completed in Excel. For the price variable, an observation will be dropped if its according price is defined as inappropriate by the PSID (if the price is 0 or 999999).

3.2. Other Supplementary Data Sources

The first supplementary dataset examined in this research is taken from the U.S. Department of Energy. This data source provides me with a wide range of miles per gallon (MPG) for different vehicles ranging from different years. Since the MPG data prior to the year 2000 is limited, the spectrum of MPG taken into this study ranges from 2000 to 2013. The dataset has three different MPG values, namely City, Highway, and combined MPG. Since the combined MPG is a weighted average based on the other two, it is the most appropriate measure for the purpose of cross-comparing different vehicles, so combined MPG will be used as the only fuel economy indicator in this study.

Unlike the PSID dataset, the data from the U.S. Department of Energy provides specific MPG for different car models within one brand. In order to match with the PSID dataset, I take an average of all the available models of vehicles in the Department of Energy dataset, given that these models are from the same year. This method can add on to the biasedness of my regression result since it is grouping normal sedan cars with SUVs and sports cars, which have very different fuel economies. However, there is no other approach since the PSID dataset only provides me with a brand-level attribute for vehicles. In the process of averaging out the car models, I separate these vehicles into two main groups that can correspond to the previously defined hybrid indicator variable. Any vehicles that are classified as an EV or an HEV are averaged together based on the brand and model year, and the non-hybrid/electric cars are averaged together similarly. After the grouping and averaging, based on years and brands, I assign to these groups the same identifier I created for the PSID dataset so that a 2010 Toyota in the Department of Energy's dataset will have the same identifying number as in the PSID dataset.

The second set of supplementary data is the GHSA dataset. This dataset supplies me with different speed limits at a state level, for both rural and urban interstates. In this research, since I am primarily focusing on the commuting time of households to and from work, I assume that the commutes happen only within urban areas, and thus only the speed limits from urban areas are considered in the study. The last set of data is the EIA dataset, which provides a set of retail prices for gasoline, in dollars per gallon, for different states in different years. Unfortunately, the data collection for the retail prices of gasoline was suspended in 2011, which is the starting focus of this study. Instead, for the year 2011 and 2013, this dataset only provides the gasoline retail prices for nine states: California, Colorado, Florida, Massachusetts, Minnesota, New York, Ohio, Texas, and Washington. Due to this, I will limit the focus of my study to these nine states only.

Thus, all of the observations where the households reside in states other than these nine will be dropped.

3.3. Summary Statistics.

Prior to analyzing the dataset, this section provides some background knowledge on the summaries of datasets. After dropping all the inapplicable observations and merging all the datasets together as well as controlling for outliers, the final pool of observations consists of 143 households in the nine states mentioned. The first variable of interest is the number of vehicles currently available in the household. In both 2011 and 2013, the number of households with two vehicles represents roughly 47% of the entire population in both years (Table 1 & 2).

When looking at the distribution of EVs and CGVs in the PSID dataset, it is evident that there is a huge discrepancy between the number of CGVs and the number of EVs. According to Table 1 and Table 2, in 2011 there is a total of 324 vehicles in 143 households; the total number of vehicles increases to 345 in 2013 for the same set of households. This overall increase in numbers of vehicles can be explained as the previous year's income of these households increases from 2011 to 2013, by approximately $30000 on average. Out of the 324 vehicles in 2011, only 3 are EVs, and in 2013, where the total number of vehicles increases to 345, there are only 4 EVs. This low number of EVs can be due to the fact that the attributes of vehicles are given only for the first three cars, so for households with more than three cars, any EVs after the third vehicle is neglected. Nonetheless, this discrepancy still exists, even when the neglected EVs are accounted for.

Table 3 illustrates the stated amount of gasoline expenses paid in dollars, both monthly and annually, by the households in 2011 and 2013. One observable change from 2011 and 2013 is that there is a slight increase in the mean of the total gasoline expenditure. On average, the same family spends approximately $10 more monthly on gasoline in 2013 than they did in 2011. After cross-checking with the EIA dataset for retail gasoline prices, which are summarized in Table 5, it can be observed that from 2011 to 2013, while there are fluctuations in retail gasoline prices from states to states, the mean retail gasoline prices from 2011 to 2013 do not change much for the nine states studied. This slight increase in gasoline expenditure can be explained by the change in travel demands. However, when looking at Table 4, the demand for travel is much higher in 2013 than it was in 2011. In 2011, the average total daily commute time for a round trip is roughly 192 minutes, or 3.2 hours. In 2013, however, the average total commute time increases to approximately 252 minutes per round trip, which is roughly 4.2 hours. The increase in standard deviation is even bigger, demonstrating that there are more families that spend more time commuting in 2013. Despite an approximately one-hour increase in commuting time on average, the gasoline expenditure summaries outlined illustrate a much lower increase in gasoline expense. Thus, it can be said that this gasoline expenditure from the PSID dataset does not necessary reflect the increase in commuting time. However, it can also be hypothesized that due to the increase in commuting habits, household gradually switch to the more fuel-efficient vehicles, and as a result, gasoline expenses are much less affected from the increasing travel demands.

Table 6 summarizes the total income of the previous year for the 143 households surveyed. Overall, there is an increase in income from 2011 to 2013, with the income in 2013 has a much higher standard deviation. This can be explained by the fact that there is a household who indicates an annual income of 3222000 for the year 2012, which has heavily influenced the distribution of income in 2013. Table 7 includes a summary for the annual insurance expenses for 2011 and 2013. The same overall trend is observed here, such that households in 2013 spend more on insurance expenses for vehicles than they did in 2011. The increased numbers of available vehicles in the dataset is a plausible explanation for this increase in annual insurance expenses for vehicles.

Table 8 demonstrates the purchasing prices for all vehicles used in this studies. Since these are the actual prices indicated by survey takers, the prices of used cars will be taken into account as well[2]. As mentioned, not all households in this population are multi-vehicles households, there are observations where the households do not own a second or third car, and in those cases the purchasing price of the second or third car is 0. This will result in an underestimated summary for the purchasing prices of the second and third vehicles. To account for this issue, I replace all the observations where the purchasing prices of the second or third vehicles to missing variables to look at the distribution of the prices. After this is recorded, these observations are changed back to 0 from being missing values, so that the vehicle utilization TCO's calculation will not be affected by missing values.

Overall, the mean purchasing prices for vehicles increases roughly $3000 from 2011 to 2013. When looking at the distributions of the years of vehicles, most of the models of vehicles in this dataset fall into the range from year 2008 to 2012. In one of their reports, the WGN News from Chicago did an average cost of cars from the year 1967. Using the data from the U.S Bureau of Economic Analysis, the WGN calculates the average cost of vehicles, with the use of the national consumer price index to account for inflations. When cross-checking with the costs of vehicles from 2008 to 2012 according to the WGN analysis, the increase in vehicles'

---

[2] Note that there are outliers in the purchasing price of the second and third vehicle in both 2011 and 2013. This outlier indicates that the prices are only $300, $400, or $500, which intuitively is very low for a vehicle. However, these outliers do not exist in the case of the first vehicle, so the regressions for efficient vehicles and EVs will not be affected.

purchasing prices are relatively similar. Based on the WGN's calculation, from 2008 to 2012, the period with the largest increase in vehicles' price is from 2009 to 2010, with an increase of $1647 (Wire, 2017). The dissimilarity in magnitude is due to the different population of vehicles in these two datasets, since there are much less types of vehicles for the vehicles of the households in the PSID dataset. However, this confirms that the increase in purchasing prices of the original PSID is in accordance with the market prices of vehicles.

4. Methodology.

One of the main focuses of this research is to determine the predicted probability of a household owning an efficient vehicle, and more specifically, an EV, as well as to find out the determinants that influence the households' vehicle purchasing decisions based on households' travel habits using OLS regressions. Similar to Fang (2008), the Probit model is also applied in this study, since the dependent variables in this study are binary, a logistic regression can also be appropriate for this study. Thus, for both efficient vehicles and EV regressions, Probit models are also implemented to compare the differences with the linear regression models. In the linear regression model, the dependent variable is considered to be continuous, so the predicted probability from the regression result can be outside the range of 0 and 1. Thus, when the dependent variable is binary, the Probit model is often preferred because it imposes a normal distribution assumption on the error term. However, as concluded by Hellevik (2007), the results of these two models are very similar. In this study, both models are implemented, and the differences in results will be discussed accordingly.

In addition to the determinants of a household owning an efficient vehicle or an EV, different TCOs will be calculated for owning efficient and non-efficient vehicles, as well as for owning EVs and CGVs. These TCOs are calculated based on the same assumption that the first listed vehicle is the primary vehicle used for commuting, as done by Borger et al. (2014). Based on these separated TCO calculations, the differences in costs of owning an efficient vehicle versus a non-efficient one, as well as an EV versus a CGV, will be distinguished, while accounting for the household's commuting habits. With regards to the TCO calculation that accounts for households' vehicles' utilizations, a further TCO analysis is done with the inclusion of the predicted probability value for a household owing an EV.

4.1. Determinants for vehicles' choices

4.1.1. Determinants for Efficient and Non-Efficient Vehicles

The first part of the study focuses on analyzing the prospects of a household owning an efficient vehicle. The calculation of efficient and non-efficient vehicle is modelled based on how the National Highway Transportation Safety Administration (NHTSA) calculates annual Corporate Average Fuel Economy (CAFE) standards using harmonic means. Different from the normal arithmetic mean, the harmonic mean can capture the fuel economy for each vehicle given that all vehicles have the same mileage driven; while the normal arithmetic mean would underestimate the total fuel used, since the arithmetic mean does not average using the total mileage driven but based on the same amount of gas (e.g. a 50 MPG vehicle would travel 50 miles and a 20 MPG vehicle would travel 20 miles).

Given the 143 vehicles available in the dataset, I apply the same calculation method using harmonic mean as the NHTSA does for CAFE standards. The harmonic mean calculation is outlined as follow:

$$\text{Mean MPG} = \frac{Total\ Number\ of\ Vehicles}{\frac{n_1}{mpg_1} + \frac{n_2}{mpg_2} + \ldots + \frac{n_n}{mpg_n}}$$

where n_1 represents the number of vehicles "1", with mpg_1 being the corresponding MPG of that vehicle. By applying this calculation, I can separate the efficient vehicles from the inefficient ones, and thus increase the number of observations for efficient vehicles, instead of just limiting my observations to EVs only.

Based on this calculation, the projected CAFE for vehicles is 19.66 MPG for 2011 and 20.28 MPG for 2013. Using these thresholds, I generate a new variable that identifies households in 2011 that own a first vehicle with an MPG higher than 19.66, and another variable that identifies households in 2013 that own a first vehicle with an MPG higher than 20.28. By doing so, the number of observations for efficient vehicles in 2011 is 75, and 86 for 2013. This increase in the numbers of efficient vehicles is in accordance with the previous observation that despite an immense increase in travel demands, the households' gasoline expenditures are not as equally influenced.

After the binary variable that separates an efficient vehicle from an inefficient one is generated, the determinants for the probability of the households owning an efficient vehicle are regressed using OLS and Probit models for 2011 and 2013. As mentioned, for the purpose of simplicity, this study uses the same assumption as Borger et al. (2014) that the first vehicle enlisted will be used as the primary vehicle, so these regressions are run based on the assumption that the first vehicle is responsible for the household's entire driving habits. Based on the studies of Bento et al. (2005), Wu et al. (2015), and Rusich and Danielis (2016), the model used in this study apply a combination of relevant variables, which includes**:** total daily average commute time in hours, purchasing price of the first vehicle, insurance expense per vehicle (calculated as the total insurance expense over the number of vehicles), total family income of the previous year, number of vehicles available, and total annual gasoline expense. The probability of a household owning an efficient vehicle is regressed as:

$$E_i = \alpha_0 + \beta_1 TotalCommute_i + \beta_2 Price_i + \beta_3 Insurance_i + \beta_4 Income_i + \beta_5 NumberVehicle_i + \beta_6 GasExpense_i + \varepsilon_i$$

where E_i indicates whether the first vehicle in household *i* is an efficient vehicle or not. A predicted value for the households owning an efficient vehicle is generated based on the regression.

The Probit model also uses the same set of independent variables as does the OLS model. From the Probit model, the estimated marginal effects are generated, which indicate the probabilities that the observed dependent variable is equal to 1. The Probit model for the probability of the household owning an efficient vehicle is outlined as follow:

$$Pr\ (E_i = 1) = \Phi\ (\beta_1 TotalCommute_i + \beta_2 Price_i + \beta_3 Insurance_i + \beta_4 Income_i + \beta_5 NumberVehicle_i + \beta_6 GasExpense_i + \varepsilon_i)$$

where $E_i = 1$ indicates that vehicle *i* is an efficient vehicle. Given the schemes of independent variables, it is hypothesized that the more the household commutes, the more likely that the household will own an efficient vehicle. Also, gasoline expense should have a negative correlation relative to the probability of owning an efficient vehicle, since intuitively, it would make sense that the more efficient vehicle would generate a lesser operating cost for households. These predictions apply for both the OLS and the Probit models.

4.1.2. Determinants for EVs and CGVs

Based on the same assumption above, I run similar linear regressions as well as Probit models to predict the probability of the first vehicle being an EV for both years. Using the hybrid indicator for the first vehicle as the dependent variable, the OLS regression for EVs uses the same independent variables as the regressions for efficient versus non-efficient vehicles. The regression model for the determinants of owning an EV is outlined as:

$$H_i = \alpha_0 + \beta_1 TotalCommute_i + \beta_2 Price_i + \beta_3 Insurance_i + \beta_4 Income_i + \beta_5 NumberVehicle_i + \beta_6 GasExpense_i + \varepsilon_i$$

where H_i is the dependent variable identifying whether the first vehicle *i* is an EV or not. Following the regression, a predicted value *yhat* for EVs is also generated to determine the probability of the first vehicle being an EV.

Similarly, the Probit model for EVs also uses the same set of independent variables. The Probit model for the probability of the household owning an EV is outlined as follow:

$$Pr\ (H_i = 1) = \Phi\ (\beta_1 TotalCommute_i + \beta_2 Price_i + \beta_3 Insurance_i + \beta_4 Income_i + \beta_5 NumberVehicle_i + \beta_6 GasExpense_i + \varepsilon_i)$$

where $H_i = 1$ indicates that vehicle *i* is an EV. Similar to the efficient vehicles regressions, it is also hypothesized in the EV's regressions that total commute time will have a positive correlation with the probability of the household owning an EV. Furthermore, based on Hidrue et al.'s (2011) observation that EVs generally cost more than CGVs, it is also hypothesized here that the higher the purchasing price of the vehicle, the higher the probability that the vehicle is an EV**.** Gasoline expense should have a negative relationship with this probability, since the EVs are considered to be more fuel efficient, so the households that own an EV would be more likely to spend less on gasoline expenditure than the households that do not own an EV.

In these regression models for both efficient vehicles and EVs, the natural log values for several variables, namely purchasing price, insurance payment, gasoline expense and total household income, will be used. The natural log values are used instead of the actual variables in order to scale down the effects of these variables into percentage changes with regards to the probability of owning an EV. The same set of regressions is run for both 2011 and 2013, given different family attributes

4.2. Total Cost of Ownership

4.2.1. Comparison between efficient and non-efficient/ EVs and CGVs

In order to calculate TCO, it is first necessary to merge all the datasets together. TCO is defined as the sum of purchasing price, cost of running the vehicle, and other maintenance costs, based on the calculations applied by previous studies (Rusich and Danielis (2015), Wu et a. (2015), Bubeck et al. et al. (2016)). Since the total average commute time from PSID is a time unit variable, this variable has to be converted into miles to calculate the cost of running the vehicle. This conversion is done by merging the data containing speed limit from the GHSA dataset with the PSID dataset based on states. An annual mileage driven variable is generated as the product of the total average commute time and its corresponding speed limit variable. As stated in the summary statistics above, the gasoline expenditure variable from the PSID dataset does not necessary reflect the change in households' travel demands, so the in the calculations of TCOs, the retail gasoline prices from a supplementary dataset is used to capture the cost based on commuting time. I thus calculate another annual gasoline cost variable based on the annual mileage driven and the average annual retail price of gasoline in the corresponding state. Using the state variable, I merge the EIA's retail prices for gasoline with the master dataset in order to match the retail price of gasoline for different states. For the vehicles' fuel economy, I combine the MPG for the matching vehicle by merging the PSID dataset with the one from the U.S. Department of Energy, using the vehicle identifier variable as the common variable. Then, the total amount of gallons of gasoline consumed is derived by dividing the annual mileage driven of a vehicle by its combined MPG, and subsequently, the product of the total amount of gallons consumed and the retail price of gasoline in that year represents the total operating cost based the total average commute time.

The TCO of a vehicle will then be the sum of the vehicle's purchasing price, its maintenance cost, and its operating cost based on household's travel demands, outlined as:

$$TCO_j = Gasprice_j * AnnualMile_j/MeanMPG_j + Price_j + Insurance_j$$

where the TCO of vehicle *j* is the sum of all the according attributes of vehicle *j* mentioned above. The mean MPG is calculated by looking at the mean of all efficient vehicles in 2011 and 2013 individually, so each year will have a different mean MPG for efficient vehicles. The same is done for non-efficient vehicles, as well as EVs and CGVs. In order to compare the different costs of vehicles in different fuel groups, the first TCOs are individually calculated for efficient vehicles and non-efficient vehicles, as well as for EVs and CGVs. The assumption that the first vehicle is the primary vehicle is still present in these TCO calculations.

4.2.2. TCO with vehicle utilization decisions.

Lastly, a TCO analysis when allowing for vehicle utilization is included. Unlike the previous two TCOs, this TCO calculation is the entire cost of households owning multiple vehicles. In this process, in addition to having the predicted probability of the first vehicle being an EV, I run separated regressions for the other two vehicles in the household individually in order to generate a weighted predicted value for each of the other vehicles being an EV. This TCO calculation does not take into account the efficient versus non-efficient vehicles analysis above due to the fact that the efficient indicator variable is generated based on the vehicles' combined MPG. In this PSID population, all households own at least one vehicle, but there are also cases when the households do not own more than one vehicle. In these cases, the MPGs for the second and third vehicle of the single-vehicle household will be 0, which will create biases if these MPGs were used to calculate the predicted value for the second, or third, vehicle being an efficient one. Unlike the efficient vehicle indicator, the EV indicator is specified from the original PSID dataset. Thus, only the predicted value for the vehicle being an EV is used in the TCO for vehicle utilization decisions.

Applying the same regression, the hybrid indicator variables for each vehicle are used as dependent variables, with the independent variable being the same as the ones for the regression above. The major difference between these regressions is the purchasing price variable, since this variable corresponds to the exact vehicle, so the price for the second vehicle is used in the regression for the second vehicle being an EV, and the price for the third vehicle is used in the regression for the third vehicle being an EV. Total daily commute time is not divided between vehicles, since there still lies an assumption that the currently regressed vehicle is used as the primary vehicle for commuting. Other than the difference in prices, the models for these regressions are identical to the first vehicle regression above.

Similarly, each regression creates a new predicted *yhat* value for the probability of the other vehicle being an EV. For each year, three separated regressions run, for the first, second, and third vehicles individually, so there will be three different predicted *yhat* values as a result. The TCO for vehicles is then calculated based on the probability of the individual vehicle regressions above. Thus, different from the previous calculations where the TCOs are only calculated for one vehicle at a time, this TCO with vehicle utilization does not apply the same assumption as before, but instead it calculates the entire TCO for the households given their choices of current vehicles and travel demands.

In order to account for all the vehicles available in the household, the three probabilities of owning an EV, as the first, second, or third car, are taken into the TCO equation. By incorporating these probabilities of owning an EV for each of the vehicles available in the household into account, utilizations between different vehicles can be accounted for. Thus, instead of having only one combined MPG for each vehicle, the total annual mileage driven variable is divided by the aggregated mean MPG for all available vehicles, with a weighted probability of owning an EV or a CGV. The TCO model is outlined as:

$$TCO_j = Gasprice_j * AnnualMile_j/WeightedMeanMPG_j + Price_j + Insurance_j$$

where $Gasprice_j$ is the annual average retail gas price of the state household j resides in, $AnnualMile_j$ is how many mile household j commutes during that year, $Price_j$ is the total price of all vehicles available in the household, insurance is the total amount of annual insurance expense paid. The WeightedMeanMPG of household j is the sum of all weighted means of MPG for both HEV and CGV, defined as:

$$WeightedMeanMPG_j = yhat_1 * MeanMPG_{hybrid} + (1 - yhat_1) * MeanMPG_{gas} + yhat_2 * MeanMPG_{hybrid} + (1 - yhat_2) * MeanMPG_{gas} + yhat_3 * MeanMPG_{hybrid} + (1 - yhat_3) * MeanMPG_{gas}$$

where $yhat_1$ represents the probability of the first car being a HEV, and $1 - yhat_1$ represents the probability of that first car being a CGV. Similarly, $yhat_2$ is the probability of the second car, and $yhat_3$ is that of the third car. MeanMPG of an EV is the average MPG of all EVs available in the households within the same year, and the same average is taken for CGV. The same model is applied for the year 2013.

5. Predicted Results and Interpretations

5.1. Probabilities Determinants Analysis

In this first set of analyses, I will only be looking at the first vehicle indicated in the dataset. Since not all families have more than one vehicle, there will be inconsistencies if all vehicles were regressed together. Thus, it is assumed that the first vehicle listed in the PSID interview is the primary vehicle, and this vehicle is responsible for the total commute time of the household. Of course, there are cases when both the head and the wife of the household travel to work simultaneously, resulting in two vehicles being utilized at the same time, but since the primary focus here is to look at the probability of a household owning an efficient vehicle, or an EV, the simultaneous utilization of vehicles within one household is neglected at this stage.

5.1.1. Probability that the first vehicle is an efficient vehicle.

This section discusses the effects of various factors on the probability of the first vehicle being efficient according to the CAFE standard calculation. According to the result, the predicted value for the first vehicle being efficient is 52.45% for 2011 and 60.14% for 2013, based on the linear regressions. The Probit models predict similar values, with 52.49% for 2011 and 61.1% for 2013.

Notably, in the regression for efficient vehicles, the number of vehicles available and the purchasing price of that vehicle both illustrate high levels of significance. In 2011, the results from both the linear and the Probit models indicate a significant relationship at the 10% level for the number of vehicles with regards to the probability of owning an efficient vehicle. In 2011, it is predicted that when the number of vehicles increases by 1, the probability of the household owning an efficient vehicle increases by 8.9% in the linear model and 9.9% in the Probit model. For 2013, this relationship displays a higher level of significance, at 5% level, and indicates a similar increase as in 2011 at 8.18% when the number of vehicles increases by 1. The Probit model also displays a similar level of significance at 5% level, with the effect being similar to that of 2011 at a 9.99% increase given the number of vehicle increases by 1.

Purchasing price is the other significant variable that influences the probability of the first car being an efficient car. Overall, in both 2011 and 2013, all models suggest an increase in the probability of owning an efficient car given an increase in purchasing price. In 2011, both the linear and Probit models indicate similar increases in probability, with 1.31% in the linear model and 1.39% in the Probit model, when the purchasing price increases by 10%. Both models display the same level of significance, 5% level, for the purchasing price variable.

In 2013, both models also indicate similar increases in probability based on purchasing prices, although with a lower level of significance. In the linear model for 2013, the results indicate a 10% level of significance for purchasing price, showing that an increase of 10% in purchasing price results in a 1.32% increase in probability of the first car being efficient. The same is observed in the Probit model, with an increase of 1.34% when the purchasing price of the vehicle increases by 10%.

The key variable in this study, which is the household's daily commute time, however, does not indicate any level of significance to the probability that the vehicle is an efficient one. In 2011, there is a positively correlated relationship between daily commute time and the probability of owning an efficient car, such that a one-hour increase in daily commute time results in an approximate increase of 1% in probability. For 2013, this relationship becomes negative, which indicates that the higher the household's travel demand is, the lower the probability that the household will own an efficient vehicle. This relationship also indicates a very small decrease, at roughly 0.2% when daily commute time increases by an hour. One plausible explanation for this change is that in 2013, due to the increase in travel demands, households gradually switch their means of transportations from driving to taking public transportations, such as buses or trains. However, as mentioned, since these relationships do not indicate any levels of significances, nothing can be said definitively regarding these coefficients.

Another interesting finding from the efficient and non-efficient vehicles comparison is the effect of gasoline expenditure on the probability of the vehicle being efficient. For both 2011 and 2013, all models suggest that there is a negatively correlated relationship between gasoline expenditure and the probability of the vehicle being efficient. Specifically, in 2011, when gasoline expenditure is doubled, the probability of the first car being efficient decreases by 0.8% in the linear model and 0.6% in the Probit model. In 2013, this decrease in probability increases to 6.9% for the OLS model and 7.5% in the Probit model, given the same increase in gasoline expenditure. Intuitively, this makes sense since the more efficient vehicle would use less gasoline, hence reducing the amount of gasoline expenditure. However, the models do not indicate any levels of significances for these effects, so it cannot be said with certainty that this negative relationship is necessarily true.

Despite being insignificant, previous year income demonstrates an interesting observation for the probability that the first car is an efficient vehicle. In both models, previous year income indicates an increase in the probability of the first vehicle being efficient in both years. It can thus be inferred from this observation that efficient vehicles are viewed as normal goods, since the probability for owning an efficient vehicle increases as income increases.

5.1.2. Probability that the first vehicle is an EV.

When applying the same method as used for the efficient vehicle's regressions above, the predicted value after regression indicates that the after-weighted average probability of owning an EV is 2.1% in 2011 from the linear regression, and 0.89% in the Probit model. That number decreases in 2013, to 1.4% in the OLS model and 0.5% in the Probit model. This decline in probability can be explained by the fact that there are more cars identified as EVs in the first vehicle group in 2011 than in 2013 (3 and 2, respectively).

Before going in depth and explaining the relationship between the independent variables and the dependent variable, it is noted that none of the relationships indicate any level of significance. This insignificance of result can be due to the issue that there is a very limited number of EVs available in the 143 households surveyed for both years. Due to this lack of observations for EVs, it is important to note that the result in this study cannot capture the entire hybrid vehicle market.

Looking at the effects each independent variable has on the probability of the first vehicle being an EV, the magnitudes are generally small for most variables. Generally, based on the direction of the signs, the trend that these independent variables have on the probability of owning an EV is consistent from 2011 to 2013, with the exception of annual gasoline expense. For 2011, annual gasoline expense has a positive relationship with the probability of owning an EV. In particular, with an increase of 100% annually in gasoline expense, the probability of the first vehicle being an EV increases by 1.09%. This means that even when the gasoline expense is doubled, the probability only increases by roughly 1%. The Probit model, on the other hand, illustrates a decrease in probability, such that the probability will decrease by 0.0017 given the same change in gasoline expenditure. Even though the two models indicate different relationships, the difference in magnitudes is negligible. However, this is not the case for 2013, where when gasoline expense is doubled, the probability of the first vehicle being an EV decreases by 1.08% for the OLS model and 0.36% for the Probit model. Nevertheless, besides the fact that these effects are small, as mentioned, none of these variables indicate any level of significances, and this can be due to the low number of EVs in the dataset.

One of the key independent variables in this regression is the total daily commute time variable. For both 2011 and 2013, both models exhibit a positive correlation between total daily commute time and the probability of owning an EV. When looking at the linear regressions, for 2011, if the total daily average commute of a household increases by one hour, the probability of the primary vehicle being an EV increases by 0.2%. This probability decreases to 0.18% in 2013. In the Probit regressions, the total daily commute time exhibits a much smaller increase in probability, with 0.08% and 0.06% for 2011 and 2013, respectively. Interestingly, there is an increase in travel demand from 2011 to 2013, and yet the probability of the primary vehicle being an EV decreases, though only by roughly 0.05%. Similar to the efficient vehicles analysis above, the same explanation, which states that households might switch to other means of transportations in response to the increasing travel demands, can be applied for the case of EVs.

Another important independent variable in this regression is the purchasing prices of vehicles. For both 2011 and 2013, purchasing prices of vehicles demonstrate a positive correlation with the probability of owning an EV, which means that the higher the price, the more likely that the car can be an EV. In 2011, when the price of the vehicle increases by 10%, the probability of that car being an EV increases by 0.179% for the linear model and 0.113% for the Probit model. This increase is even smaller in 2013, being 0.037% in the linear model and 0.0028% in the Probit model, given that the price also increases by 10%. From 2011 to 2013, it is observed here that given the same percentage increase in purchasing price, the probability of the vehicle being an EV decreases. A plausible explanation for this observation is that the prices of EVs increase from 2011 to 2013, resulting in a decrease in the probability of being an EV when purchasing price increases.

Previous year income indicates a negative relationship to having an EV as the primary vehicle for both 2011 and 2013. In 2011, given that the previous year's income is doubled, the probability of owning an EV decreases by 3% in the linear model and 1.43% in the Probit model; in 2013, this probability decreases by much less, at 0.98% for the OLS model and 0.36% for the Probit model, given the same increase in previous year's income. Thus, one interpretation that can be inferred based on this result is that for both 2011 and 2013, EVs are viewed as inferior goods given that the probability of owning an EV decreases when income increases. However, another interpretation is that, with regards to income, the probability of owning an EV as the primary vehicle is experiencing a decrease at a decreasing rate in terms of magnitude. Since the probability indicates a negative relationship, a decrease in magnitude is in fact an increase in probability of owning an EV over years. As illustrated in Table 6, there is an overall increase of approximately $30000 in annual income from 2010 to 2012 for the average household. An observable trend here is that given this increase in income, the probability of owning an EV with regards to income increases as well. If the same trend persists, it can be anticipated that the probability of owning an EV will keep on increasing, and eventually will become positive, holding the change in income constant. This interpretation is in contrast with the previous one, which infers that EVs are inferior goods. A conclusion can be drawn from these results that EVs are not necessarily inferior goods, but the reasons why households are deterred from owning an EV can be due to its current limitations, such as the high purchasing price, range limitations, etc. Yet, given the cross-comparison between 2011 and 2013, it is observable that there is an increasing acceptance toward EVs as income increases.

The number of available vehicles in the household displays an overall positive trend on the probability of owning an EV. In 2011, if the number of vehicles in an average household increases by 1 unit, the probability of the first car being an EV increases by 0.09% in the linear model and 0.22% in the Probit model. In 2013, this probability increases by 0.15% for the linear model, but decreases by 0.03% in the Probit model, providing there is one more vehicle in the household. Intuitively, the household would only be able to afford an additional vehicle when the household's income increases. This also corresponds to the observation above that annual income for these 143 households increases from 2011 to 2013, which can result in a higher possibility in affording an additional vehicle. Hence, based on these regressions' results, given that the probability of owning an EV increases from 2011 to 2013 with regards to the number of available vehicles, this supports the inference that there is an increasing acceptance toward EVs.

Lastly, insurance expense, which is the variable that represents the vehicle's maintenance cost, illustrates a positive correlation with the probability of owning an EV as the primary car. However, from 2011 to 2013, the effects of insurance expenses on the probability of owning an EV decreases, from 0.2% to 0.1% given an increase of 10% in insurance expense. As noticed before, the predicted probability of owning an EV is lower in 2013 than in 2011, based on both the OLS and the Probit models. Thus, a causal effect is observed here, such that when there are less EVs available, the effects of insurance expenses associated with EVs decreases as well. Once again, none of the variables aforementioned indicate a level of significance, even at the 10% level, with regards to the probability of owning an EV as the first car. The result here is then inadequate to completely reflect the determinants for households for choosing an EV as a replacement for the primary vehicle.

5.2. Total Cost of Ownership Analysis

In the following set of analyses, the TCO calculations for different types of vehicles with different levels of fuel efficiencies are discussed. The first two analyses focus on the TCO calculations where the assumption that the first vehicle is the primary vehicle still exists. The last TCO analysis takes into account the households' vehicle utilization decisions, and thus will illustrate the entire TCO for all available vehicles in the household. For this TCO analysis, in order to see if there are any differences when vehicle utilization is not allowed for, there is an additional TCO calculation that only uses the normal mean MPG for all vehicles instead of the weighted mean MPG discussed above.

5.2.1. Total Cost of Ownership for Efficient vehicles and Non-efficient vehicles

Table 14 demonstrates the summaries of TCOs for the efficient and non-efficient vehicles in 2011 and 2013. In both years, it is observed here that the efficient vehicles are generally more expensive to own. This can be explained by the observation that the prices of the efficient vehicles are generally more expensive than that of the non-efficient vehicles, as illustrated in table 8. Specifically, in 2011, it costs almost $9000 more on average to own an efficient vehicle. In the same year, the price difference between the efficient and non-efficient vehicles is roughly $7000 (Table 9). The gap in TCOs in 2013 is much smaller than in 2011, with the efficient vehicle costing only $2000 more to own than the non-efficient vehicle. The purchasing price gap for efficient and non-efficient vehicle is also much smaller in 2013, with the efficient vehicle costing only $4800 more to buy. However, the increase in TCOs from 2011 to 2013 is much greater for the non-efficient vehicles than for the efficient ones. This increase can be partially explained by the much higher increase in purchasing prices for non-efficient vehicles in 2013. Another explanation for this increase is the increase in travel demands of households from 2011 to 2013. Intuitively, as travel demands increase, it will cost much more to own a non-efficient vehicle since the efficient vehicle has a higher level of fuel efficiency, so the higher the travel demands, the more benefits can be enjoyed from owning an efficient car. As expected, it can be observed here that as travel demands increase, the marginal increase in TCO for the non-efficient vehicle is much more than that for the efficient one. Notice here that in 2013, the gap in TCOs is not in accordance with the gap in vehicles' purchasing prices. One explanation for this is that households who already owned an efficient vehicle in 2011 do not need to spend more money on purchasing an additional efficient vehicle, and hence these households enjoy more benefits, or incur less costs, from already owning an efficient vehicle.

In order to verify for the explanation above, I further look into a 5-year interval TCO analysis. Since the above TCOs' calculations are only the costs of owning the vehicle for the current year, households who already owned an efficient vehicle will not have enough time to recover the vehicles' purchasing costs from just one year of driving. In this 5-year interval TCO analysis, it is assumed that the households' travel demands as well as gasoline prices do not change from the base year. Since there are fluctuations in travel demands from 2011 to 2013, there are two 5-year interval TCO calculations, with one using 2011 as the base year and the other using 2013. These 5-year interval TCOs are calculated by multiplying the current variable costs, which are operating cost and maintenance cost by 5, while the purchasing cost of the vehicle remains constant.

When using 2011 as the base year, the efficient car still costs more to own, but the gap is much smaller. In a 5-year period, on average it costs approximately $16000, or $3200 a year, more to own an efficient vehicle than to own a non-efficient one. However, this cost difference is reversed in 2013. When 2013 is used as the base year, the efficient vehicle costs less to own than the inefficient one. On average, it costs almost $13000, or $2600 a year, less to own an efficient car. Thus, both observations here verify the above interpretation that after years of using, households will gradually incur less in total costs when owning an efficient vehicle than a non-efficient vehicle, confirming that the marginal cost of owning an efficient vehicle is less when compared that of owning a non-efficient vehicle.

5.2.2. Total Cost of Ownership Comparisons between EVs and CGVs.

A similar analysis is implemented to study differences in TCOs for EVs and CGVs. In 2011, on average, it costs approximately $3000 more to own an EV than to own a CGV. However, the TCO for CGVs has a higher standard deviation than that of EVs, meaning that the TCO for CGVs fluctuates much more than TCO for EVs does. In 2013, the gap in TCOs for EVs and CGVs is reversed, with EVs costing approximately $8000 less on average to own. Similar to the TCOs for efficient and non-efficient vehicles above, the TCOs for EVs and CGVs here are only limited to the cost of a 1-year period. Thus, in order to look for the potential reduced operating costs, the same method as done for the TCOs of efficient and non-efficient vehicles is applied, so 5-year interval TCOs are generated for EVs and CGVs as well.

Using 2011 as the base year, the 5-year period TCO indicates that EVs cost $10000, which is $2000 a year, more to own on average. This gap is very similar to the 1-year period TCO calculated above, demonstrating that the costs of owing EVs do not decrease by much even after five years of commuting. Among the previous studies that calculate TCOs for EVs and CGVs, Rusich and Danielis's (2015) is the other study that also looks at a 5-year interval TCO for EVs and CGVs. Their results indicate that even after a 5-year period, EVs are still more expensive to own than CGVs, which is similar to the TCO results in this study when 2011 is used as the base year (Rusich and Danielis, 2015). However, despite being a small decrease in TCO's difference (from $3000 to $2000 a year), this observation still demonstrates that there is a diminishing marginal cost in owning an EV, but a 5-year period is not long enough for an EV to recover its higher initial cost[3]. This is similar to Thiel et al.'s (2010) results, which state that the payback period for EVs in general is at least 20 years when using 2010 as the base year.

When using 2013 as the base year, the gap in 5-year TCOs shows that EVs cost roughly $16000 less to own than CGVs cost over a 5-year period. This can be translated to a $3200 less in TCO per year for EVs, which is much less than in the 1-year period comparison, where it costs $8000 less to own an EV. This does not make sense based on the previous observation, which indicates that a 5-year period from 2011 is not long enough for EVs to be less costly than CGVs. However, when looking at the distribution of purchasing prices for EVs and CGVs from Table 9, this gap in TCOs can be explained. It is observed here that in 2013, an CGV costs more to own than an EV does on average. This does not make sense according to Hidrue et al.'s (2011) as well as Rusich and Danielis's (2015) observations that EVs generally cost more than CGVs. However, in this case, since vehicles with low MPG but high purchasing prices, such as sport cars or SUVs, are also included in the group of CGVs, resulting in an overestimated purchasing price for CGVs. Moreover, the low number of observations for EVs can also be another reason that affects the results for TCOs of EVs, since a very small number of EVs is used in this calculation.

Despite being overestimated, the TCO when using 2013 is most accurately compared to Wu et al.'s (2015) results since the base year used in their study is 2014. Since Wu et al. (2015) look at the TCO/km of vehicles in Germany, a TCO/mile variable is derived by having the TCO divided by annual mileage driven in 2013 to more accurately compare with Wu et al.'s (2015) results. Also, since Wu et al.'s (2015) study is conducted in Germany, the units used are EUR and kilometers, so in order to more accurately assess the results, the TCO in this study is

[3] Table 9 indicates the higher purchasing costs for EVs in 2011.

converted into EUR/km terms using the 2013 EUR to USD exchange rate (1 EUR = 1.33 USD)[4][5]. Wu et al.'s (2015) result in the medium range suggests that CGVs have the lowest TCO as of 2014, at 42 EUR cents/km, as compared to 44 EUR cents/km for HEVs, or roughly 49 EUR cents/km for EVs when all EVs models (HEVs, PHEVs, and BEVs) are averaged. However, the same is not seen in the PSID dataset, as the 2013 base year TCO model suggests that EVs cost less to own, at roughly 69 USD cents/mile, or 58 EUR cents/km after conversion, but this is still relatively similar to Wu et al.'s (2015) observation. The post-conversion TCO for CGVs is much higher, at 1.9 EUR/km. Regardless of the difference in magnitudes, the two results suggest different indications, with one stating that CGVs have lower TCO and the other stating otherwise. When looking at the actual TCOs for different vehicles, even though the TCOs for EVs are relatively similar with only a 10 EUR cents/km gap, the TCOs for CGVs from the two results are very different. Again, as mentioned, since the PSID dataset does not separate SUVs and sport cars from normal sedan vehicles, the TCO for CGVs in general is largely overestimated, which can be the reason why the TCO for CGVs in this study is much higher than that of Wu et al.'s (2015). Furthermore, due to the low numbers of EVs available, the result from this study can be largely affected as well.

5.2.3. Total Cost of Ownership for all Vehicles in 2011 and 2013

As mentioned in the methodology section, the TCO computation is formulated based on all the probabilities of each of the three vehicles being an EV. Thus, this method will allow for vehicle utilization between different vehicles available in the household to be considered in the TCO model.

[4] The exchange rate is taken from xe.com.
[5] 1 mile is equal to 1.6 kilometers.

Table 17 demonstrates the summary statistics for TCOs of all vehicles in both 2011 and 2013. Overall, there is an increase of more than $6000 in TCO from 2011 to 2013 for the 143 surveyed households. There are several factors that contribute to this increase. First, for the same set of households, there are 21 more vehicles in the dataset for 2013 than for 2011. As more vehicles are purchased, the total purchasing price for vehicles of a household increases as well. Thus, this increase in the number of vehicles can directly lead to an increase in the purchasing prices of vehicles, which eventually results in an increase in TCO. Another plausible explanation for this increase in TCO is due to the increase in daily travel demand. Since the TCO model does not take into account gasoline expense but uses actual gasoline retail prices based on states, the cost of running vehicles is entirely dependent on the amount of time spent in commuting. Thus, despite the relatively insignificant increase in gasoline expenditure in the PSID dataset, the gasoline cost variable used in the TCO is directly correlated to the daily commute time variable. Given that there is an increase of roughly one hour in total daily commute time, as illustrated in table 4, the gasoline cost variable can not only precisely reflect this increase, but also more precisely illustrate the amount of gasoline spent on commuting using vehicles.

In order to look at the difference in TCO when vehicle utilization decisions are included, a calculation of TCO without vehicle utilization decisions is included for both years. For this TCO calculation, instead of using the weighted mean MPG derived from the predicted values, the normal mean MPG from available vehicles is used. The results show that when vehicle utilization is included, the TCOs in both years are much less costly. In 2011, when vehicle utilization is included, the TCO for all vehicles is about $7000 less. That gap is even bigger in 2013, with a $9000 decrease in TCO when vehicle utilization is included. As Spiller's (2012) conclusion states that by not allowing for utilization between bundles of vehicles, past studies have underestimated the elasticity of demand for gasoline, which means that in reality consumers are much more sensitive to changes in gasoline prices when the vehicle utilization option is allowed. My results here show that the TCO for vehicles is overestimated without the inclusion of vehicle utilization. When the option to utilize the households' bundles of vehicles is allowed, households would make their commuting as well as vehicle purchasing decisions accordingly in order to minimize their costs. Thus, similar to Spiller's (2012) results, my results also indicate that when allowing for vehicle utilization, households are much more sensitive to costs and will utilize their bundles of vehicles in such a way that can minimize their costs.

6. Concluding Remarks and Policy Implications

As the regression result suggested, given the current dataset, there are not enough EVs in the population for the results to indicate a high level of significance. Yet, it is noteworthy to point out the method and model employed in this research, which has not been done by previous studies. Thus, given a better dataset with more observations for EVs, it is still worthwhile to implement the same method used in this study, despite the current level of insignificance.

Based on the TCOs results, it is evident that while the efficient vehicles might cost more to own at the current period, as the efficient vehicles are utilized more often and over a longer period of time, they will incur less costs and gradually generate more benefits. The same effect can be observed for EVs, although with a much slower rate of returns. Thus, the decision whether to own an efficient vehicle or an EV narrows down to how much the user expects to utilize the vehicle. In order to optimally utilize one's vehicle choice decision, two main questions should be asked: how often the user intends to drive the vehicle, and how long the user intends to keep the vehicle for.

Overall, this study provides two important insights. First, even when the TCO for efficient vehicles is seemingly cheaper than that of non-efficient vehicles for the current year, when looking at a longer period of time, the extra benefits from having higher fuel-economies can gradually payback for the extra costs generated from the initial higher purchasing prices. Second, when the option of vehicle utilization is allowed, households have more freedom to move between their alternatives, and thus able to minimize their total costs when utilizing vehicles. Based on these insights, some policy implications can be drawn. First, in order to promote the uses of more fuel-efficient vehicles, governments can introduce different types of subsidies to decrease the initial high purchasing cost, and thus give consumers more incentives to buy fuel-efficient vehicles. Another implication from this study is that instead of just showing the levels of fuel-efficient of vehicles to consumers, consumers should be educated on how to make vehicles purchasing's decisions that can best fit their travel demands.

The levels of insignificance in this study can be due to certain limitations of the dataset used, and some additional regressions can be tested to better understand the costs and benefits of different levels of fuel efficiencies for different vehicles. As stated earlier, since the PSID dataset does not separate different types of CGVs, such as SUVs and sport cars from normal sedans, the estimations for CGVs could have been largely influenced. A flaw of this study is the low number of observations for EVs, which prevents the study from accurately capturing the entire EV market. Thus, it is suggested based on this study that even though it offers a wide range of household's characteristics, the PSID dataset is not necessarily suitable for transportation research due to the inadequacy of the relevant variables. Another flaw of the study is the assumption that drivers travel consistently at the speed limit in their daily commutes. This is not accurate realistically since the speed at which the vehicles are driven always fluctuates. Thus, due to this assumption, the results of TCOs here may have been overestimated when compared to actual TCOs.

Other than the limitations of this study, there are also other unaccounted factors in this study that further research can explore. First, even though this study looks at the TCO for EVs, the cost of electricity was not accounted for. While the PSID dataset does offer a variable that summarizes the electricity cost of a household, this cost is the entire electricity cost for the household, which includes the costs generated from many other electrical appliances, so it will be inaccurate to include this cost for EVs' operating costs. Thus, this cost should also be accounted for, in order to more accurately assess the benefits and costs of EVs. Another important aspect that was not included in this analysis is the externalities of EVs. These externalities can either be costs or benefits, which can either incentivize or deter users from choosing an EV. For instance, the reduced level of CO_2 emissions is an external benefit that may incentivize some users to buy an EV, but the limitation in driving range is an external cost that can discourage users from buying. Although these externalities can hold different importance for different users, and it is difficult to measure how environmentally concern a user is or how inconvenient a user feels from the limited range problem, these are still important aspects of EVs that should be addressed since they are what separate EVs from CGVs.

Appendix

Table 1: Numbers of Vehicles in 2011

Numbers of Vehicles	Numbers of Households
1	30
2	68
3	31
4	7
5	6
6	1
Total Observation	143
Total Numbers of Vehicles	323
Total Numbers Of Hybrids (within the first three vehicles)	3

Table 2: Numbers of Vehicles in 2013

Numbers of Vehicles	Numbers of Households
1	26
2	67
3	29
4	12
5	6
6	2
7	1
Total Observation	143
Total Numbers of Vehicles	344
Total Numbers Of Hybrids (within the first three vehicles)	4

Table 3: Gasoline Expense (in Dollars)

Variable	Observations	Mean	Std. Dev.	Min	Max
Gasoline Expense in 2011	143	337.8112	286.5128	30	2000
Gasoline Expense in 2013	143	348.5245	291.9032	50	2000

Table 4: Total Average Daily Commute Time of the Household (in minutes)

Variable	Observations	Mean	Std. Dev.	Min	Max
Daily Commute Time 2011	143	192.2378	339.5229	1	1200
Daily Commute Time 2013	143	252.0769	391.5729	1	1200

Table 5: Gasoline Retail Prices in Different State

Retail Gas Price	***2011***	***2013***
California	3.863	3.933
Colorado	3.446	3.47
Florida	3.55	3.572
Massachusetts	3.592	3.627
Minnesota	3.55	3.496
New York	3.804	3.837
Ohio	3.505	3.506
Texas	3.429	3.388
Washington	3.768	3.691
Mean	3.5495	3.5486
Standard Deviation	.1505	.1810

Table 6: Previous Year's Income Summary

Variable	Observations	Mean	Std. Dev.	Min	Max
Income 2010	143	105398.4	120399.5	7976	885000
Income 2012	143	137948.6	296542	7200	3222000

Table 7: Insurance Expense Summary

Variable	Observations	Mean	Std. Dev.	Min	Max
Insurance 2011	143	2179.545	1427.492	90	7800
Insurance 2013	143	2506.629	2442.005	12	18000

Table 8: Purchasing Prices of Vehicles in 2011 and 2013

Variable	Observations	Mean	Std. Dev.	Min	Max
Purchasing Price of Vehicle 1 - 2011	143	18887.54	13168.37	1000	82000
Purchasing Price of Vehicle 2 - 2011	143	16505.57	16430.91	300	55000
Purchasing Price of Vehicle 3 - 2011	143	4005	4113.07	400	16000
Purchasing Price of Vehicle 1 - 2013	143	21911.64	13795.92	1200	65000
Purchasing Price of Vehicle 2 - 2013	143	17462.66	17795.1	500	75000
Purchasing Price of Vehicle 3 - 2013	143	7564.286	6756.946	1500	22000

Table 9: Purchasing Prices of Vehicles in 2011 and 2013 – Grouped into types

Variable	*Observations*	*Mean*	*Std. Dev.*	*Min*	*Max*
2011 – Efficient Vehicle	75	22048	11102.65	1000	82000
2011 – Non-Efficient Vehicle	68	15401	12602.7	1200	55000
2013 – Efficient Vehicle	86	23827	10782.47	1200	65000
2013 – Non-Efficient Vehicle	57	19021	14980.31	2300	62000
2011 – Hybrid Vehicle	3	20000	10000	10000	30000
2011 – Gasoline Vehicle	140	18863.7	12330.95	1000	82000
2013 – Hybrid Vehicle	2	16500	4949.747	13000	20000
2013 – Gasoline Vehicle	141	21988.4	12403.3	1200	65000

Table 10: Regressions for Efficient – Non-efficient Vehicles in 2011

VARIABLES	Linear Efficient_2011	Probit Efficient_2011
Daily Commute Time	0.00913	0.0103
	(0.00625)	(0.00807)
Previous Year Income (log)	0.0301	0.0368
	(0.0697)	(0.0670)
Number of Vehicles	0.0887*	0.0994*
	(0.0459)	(0.0520)
Vehicle Insurance (log)	0.0747	0.0850
	(0.0662)	(0.0791)
Annual Gasoline Expense (log)	-0.00809	-0.00605
	(0.0588)	(0.0678)
Purchasing Price (log)	0.131**	0.139**
	(0.0630)	(0.0586)
Constant	-1.738**	
	(0.719)	
Predicted Value	0.52	0.52
Observations	143	143
R-squared	0.144	

Table 11: Regressions for Efficient – Non-efficient Vehicles in 2013

VARIABLES	Linear Efficient_2013	Probit Efficient_2013
Daily Commute Time	-0.00273	-0.00318
	(0.00609)	(0.00650)
Previous Year Income (log)	0.0606	0.0594
	(0.0629)	(0.0685)
Number of Vehicles	0.0818**	0.100**
	(0.0326)	(0.0438)
Vehicle Insurance (log)	0.00455	0.00650
	(0.0535)	(0.0584)
Annual Gasoline Expense (log)	-0.0693	-0.0759
	(0.0651)	(0.0700)
Purchasing Price (log)	0.132*	0.134*
	(0.0668)	(0.0718)
Constant	-1.021	
	(0.765)	
Predicted Value	0.6	0.61
Observations	143	143
R-squared	0.119	

Table 12: Regressions for EVs - CGVs in 2011

VARIABLES	Linear Hybridcar_2011	Probit Hybridcar_2011
Daily Commute Time	0.00206	0.000787
	(0.00291)	(0.00104)
Previous Year Income (log)	-0.0300	-0.0143
	(0.0283)	(0.0126)
Number of Vehicles	0.000942	0.00222
	(0.00741)	(0.00961)
Vehicle Insurance (log)	0.0215	0.0158
	(0.0169)	(0.0140)
Annual Gasoline Expense (log)	0.0109	-1.68e-05
	(0.0350)	(0.00885)
Purchasing Price (log)	0.0179	0.0113
	(0.0161)	(0.0123)
Constant	-0.0552	
	(0.196)	
Predicted Value	0.021	0.0089
Observations	143	143
R-squared	0.035	

Table 13: Regressions for EVs - CGVs in 2013

VARIABLES	Linear Hybridcar_2013	Probit Hybridcar_2013
Daily Commute Time	0.00179	0.000611
	(0.00221)	(0.000769)
Previous Year Income (log)	-0.00980	-0.00364
	(0.0142)	(0.00746)
Number of Vehicles	0.00154	-0.000294
	(0.00438)	(0.00587)
Vehicle Insurance (log)	0.0150	0.0116
	(0.0112)	(0.0112)
Annual Gasoline Expense (log)	-0.0108	-0.00364
	(0.0122)	(0.00828)
Purchasing Price (log)	0.00367	0.000280
	(0.0107)	(0.00721)
Constant	0.0628	
	(0.107)	
Predicted Value	0.014	0.0049
Observations	143	143
R-squared	0.023	

Table 14: TCO for efficient and Non-efficient vehicles

Variable	Observations	Mean	Std. Dev.	Min	Max
TCO2011_Efficient	75	34710.19	22155.92	2690.732	97164.54
TCO2011_Non-efficient	68	26190.48	21292.85	3831.924	81106.33
TCO2013_Efficient	86	37237.94	22933.67	6246.186	90319.15
TCO2013_Non-efficient	57	35901.86	29833.77	3499.322	110179

Table 15: TCO for efficient and Non-efficient vehicles – 5-year interval

Variable	Observations	Mean	Std. Dev.	Min	Max
TCO2011_Efficient	75	85357.75	95369.85	8761.349	359156
TCO2011_Non-efficient	68	69346.8	95443.46	8768.802	386258
TCO2013_Efficient	86	90879.65	97552.05	11230.93	337648.1
TCO2013_Non-efficient	57	103425.2	124990.3	8296.609	412508.8

Table 16: TCO for Gasoline and Hybrid Vehicles

Variable	Observations	Mean	Std. Dev.	Min	Max
TCO2011_Hybrid	3	34774.72	19749.06	14146.02	53507.39
TCO2011_Gas	140	31818.47	22850.5	2755.893	105833.8
TCO2013_Hybrid	2	30377.92	18943.77	16982.64	43773.19
TCO2013_Gas	141	38101.52	26079.84	3466.842	103954

Table 17: TCO for Gasoline and Hybrid Vehicles– 5-year interval

Variable	Observations	Mean	Std. Dev.	Min	Max
TCO2011_Hybrid	3	93873.58	82738.7	30730.11	187536.9
TCO2011_Gas	140	83637.54	97643.86	8359.639	402502.3
TCO2013_Hybrid	2	85889.57	74919.87	32913.23	138865.9
TCO2013_Gas	141	102554	109527.6	8134.208	401070.1
TCO2013_Hybrid (EUR/km)	2	0.5798989	0.640526	0.1269786	1.032819
TCO2013_Gas (EUR/km)	141	1.979646	3.606056	0.1549125	34.62095

Table 18: TCO for all vehicles when vehicle utilization is allowed.

Variable	Observations	Mean	Std. Dev.	Min	Max
TCO_2011	143	29297.96	21130.58	254.255	112693.8
TCO_2013	143	35430.94	23975.75	2611.435	125491.1
TCO_2011 (without vehicle utilization)	143	36506.08	27986.47	2746.861	152792.3
TCO_2013 (without vehicle utilization)	143	44425.4	31295.67	2762.409	136201.1

Bibliography:

Alternative Fuels Data Center (AFDC) (2016). *Alternative Fuels Data Center: Hybrid and Plug-In Electric Vehicles*. <*http://Afdc.*energy.gov.>. Retrieved on April 2nd,2017.

Al-Alawi, B. & Bradley, T. (2013). Total cost of ownership, payback, and consumer preference modeling of plug-in hybrid electric vehicles. *Applied Energy*, *103*, 488-506.

Bento, A., Cropper, M., Mobarak, A., & Vinha, K. (2005). The Effects of Urban Spatial Structure on Travel Demand in the United States. *Review Of Economics And Statistics*, *87*(3), 466-478.

Bhat, C. (2005). A multiple discrete–continuous extreme value model: formulation and application to discretionary time-use decisions. *Transportation Research Part B: Methodological*, *39*(8), 679-707.

Bubeck et al., S., Tomaschek, J., &Fahl, U. (2016). Perspectives of electric mobility: Total cost of ownership of electric vehicles in Germany. *Transport Policy*, *50*, 63-77.

De Borger, B., Mulalic, I., &Rouwendal, J. (2016). Substitution between cars within the household. *Transportation Research Part A: Policy And Practice*, *85*, 135-156.

Derakhshan, A., Khademi, A., Khademi, S., Yusof, N., & Lee, M. (2015). Review of Discrete-continuous Models in Energy and Transportation. *Procedia CIRP*, *26*, 281-286.

Fang, H. (2008). A discrete–continuous model of households' vehicle choice and usage, with an application to the effects of residential density. *Transportation Research Part B: Methodological*, *42*(9), 736-758.

Gilmore, E. & Lave, L. (2013). Comparing resale prices and total cost of ownership for gasoline, hybrid and diesel passenger cars and trucks. *Transport Policy*, *27*, 200-208.

Hellevik, O. (2007). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity*, *43*(1), 59-74.

Hidrue, M., Parsons, G., Kempton, W., & Gardner, M. (2011). Willingness to pay for electric vehicles and their attributes. *Resource And Energy Economics*, *33*(3), 686-705.

Kihm, A., & Trommer, S. (2014). The new car market for electric vehicles and the potential for fuel substitution. *Energy Policy*, *73*, 147-157.

Li, Z., Jiang, S., Dong, J., Wang, S., Ming, Z., & Li, L. (2016). Battery capacity design for electric vehicles considering the diversity of daily vehicles miles traveled. *Transportation Research Part C: Emerging Technologies*, *72*, 272-282.

McDonald, J., & Moffitt, R. (1980). The uses of Tobit analysis. *The Review Of Economics and Statistics, 62(2),* 318.

Panel Study of Income Dynamics, public use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI. <https://psidonline.isr.umich.edu>. (Retrieved on March 24th, 2017).

Rusich, A. &Danielis, R. (2015). Total cost of ownership, social lifecycle cost and energy consumption of various automotive technologies in Italy. *Research In Transportation Economics*, *50*, 3-16.

Speed Limit Laws. *Governors Highway Safety Association (GHSA)*. <http://www.ghsa.org>. (Retrieved on March 30th, 2017).

Spiller, E. (2012). Household vehicle bundle choice and gasoline demand: A discrete-continuous approach. *SSRN Electronic Journal*.

Tamor, M. &Milačić, M. (2015). Electric vehicles in multi-vehicle households. *Transportation Research Part C: Emerging Technologies*, *56*, 52-60.

Tseng, H., Wu, J., & Liu, X. (2013). Affordability of electric vehicles for a sustainable transport system: An economic and environmental analysis. *Energy Policy*, *61*, 441-447.

Thiel, C., Perujo, A., & Mercier, A. (2010). Cost and CO2 aspects of future vehicle options in Europe under new energy policy scenarios. *Energy Policy*, *38*(11), 7142-7151.

U.S. Department of Energy. *Energy efficiency & renewable energy.* Fuel Economy Data. <http://www.fueleconomy.gov>. (Retrieved on March 30th, 2017).

U.S. Retail Gasoline Prices. *Energy Information Administration (EIA) - Official Energy Statistics from the U.S. Government*. <https://www.eia.gov>. (Retrieved on March 29th, 2017).

Wire, T. (2017). *The average car now costs $25,499 – how much was a car the year you were born?. WGN-TV. Retrieved from* <http://wgntv.com>. (Retrieved on April 28th,2017).

Wu, G., Inderbitzin, A., &Bening, C. (2015). Total cost of ownership of electric vehicles compared to conventional vehicles: A probabilistic analysis and projection across market segments. *Energy Policy*, *80*, 196-214.

Wu, X., Aviquzzaman, M., & Lin, Z. (2015). Analysis of plug-in hybrid electric vehicles' utility factors using GPS-based longitudinal travel data. *Transportation Research Part C: Emerging Technologies*, *57*, 1-12.