1. Background

The environmental sustainability has grasped tremendous attention of the government in recent years and different policies have been proposed and implemented to address this issue. One of the most widespread policies to foster environmental sustainability is the government’s supports for the purchases of hybrid vehicles. The Energy Policy Act of 2005 provides tax credits amounting to $3000 at most for consumers who purchase energy efficient vehicles.

The tax credit provided for purchasers of hybrid vehicles is considered as a powerful tool to promote the hybrid vehicles, which typically achieve more efficient fuel economy and lower emissions. Indeed, ever since the introduction of the tax credits in 2006, sales increased dramatically from less than 10,000 in 2000 to over 400,000 in 2014.

However, compared to the overall demand for autos in the US, the market for hybrid vehicles is still a niche. For example, the hybrid vehicles take account for no more than 2.5% of the entire auto sales in December 2014, and represent merely 15% of the whole market in the whole year 2014. Despite a booming market for autos as a result of economic recovery, the sales for hybrid vehicles are decreasing due to recent drop in gas prices.

The question that we aim to answer is to what extent is the potential consumers’ decisions to purchase hybrid vehicles are determined by the tax credits and fluctuating gas prices. This will not only help us understand the rationale behind the individual choices, but will also guide the policy toward environmental sustainability: Should the government provide more benefits to stimulate the demand the hybrid cars when the gas price is low? Should the government take measures to stabilize the gas price?

We start by articulating a theoretical framework of consumers’ individual choice, which is affected by the vehicle characteristics (including MPG, vehicle size, hybrid or plug-in supported among others), consumer attributes (including incomes and demographics), gas prices and tax credits for hybrid vehicles. We then aggregate the consumers’ individual choice among different auto models into the market levels. By matching the market share with the observed sales data, we extract the consumers’ preference for hybrid vehicles.

While the theory based on consumers’ choice at the micro level provides a framework with which to analyze the effect of tax credits on social welfare, whether such policies foster environmental sustainability in the ever changing external conditions (such as fluctuating gas...
prices) remains an empirical question. The data of auto sales are retrieved by implementing web-scraping programs. Our dataset is cross-examined from multiple sources. The sales data are then combined with the auto characteristics and consumer attributes. The estimated is then conducted according to the theoretical framework articulated above.

To evaluate the effect of tax credits on environmental sustainability, we measure the environmental externality from non-hybrid vehicles by the CO2 emission. We will perform counterfactual analysis -- that is, how many more tons of CO2 would be emitted if the policy providing tax credits was not implemented. We will also provide predictions on the policy implication at different gas prices -- how many tons of CO2 would be emitted if the tax credits policy is implemented and what the optimal level of the credit is.

2. Data

Our dataset covers the sales of vehicle in the US over the time span starting from January 2010 to December 2014. We include a total of 270 vehicle models from 34 brands. These data are retrieved from www.goodcarbadcar.net with a web-scraping program. The data are then cross-examined with the annual report from www.statista.com and the National Automobile Dealers Association (www.nada.org).

The next figure illustrates the time trend of auto sales from 2010 to 2014. The vertical axis represents the number of autos purchased and the horizontal axis represents the time in step of months. The grey area denotes the sales of non-hybrid vehicles and the black area denotes the sales of hybrid vehicles. There are several empirical patterns that can be observed from inspecting the figure. Firstly, the total sales of vehicles are gradually increasing from nearly 600,000 in the January of 2010 to over 1,200,000 in the January of 2014, which means that the market size almost doubles. Secondly, the auto sales exhibit a strong seasonal pattern where the market expands during the spring and shrinks during the summer. Such a within-year variation indicates that the seasonal effect plays an important role in consumers’ decision to purchase vehicles. Thirdly, despite the market expansion of the industry as a whole, the sales of hybrid car remains stable at 200,000 per month. That is, although the demand for automobiles is stimulated during the years, there is not much increase in the purchases of hybrid vehicles.

One reason that we do not observe increase in the demand for the hybrid vehicles might be that the gas prices remains roughly stable during the time span and the consumers expect declines in gas prices. To verify this hypothesis, we need to introspect consumers’ decision process. In the next section, we describe a model that aggregates consumers’ individual choice into the market shares.
3. Theoretical Model

To estimate the demand for hybrid vehicles, we base the theoretical model on the potential consumers’ individual utility function, which is determined primarily by the characteristics of the auto. The goal of the estimation is to estimate the consumers’ taste for auto attributes such as MPG. The consumers’ decision to purchase vehicles is modeled as the random coefficient demand where consumers vary in their tastes for different characteristics of autos.

A representative consumer $i$’s utility from purchasing vehicle $j$: 

$$u_{ij} = \alpha_i p_j + \sum_k x_{jk} \cdot \beta_{ijk} + \epsilon_{ij},$$

where $p_j$ is the price for vehicle $j$ and $(x_{j1}, \ldots, x_{jK})$ is a vector of vehicle characteristics. The consumer’s taste for vehicle is described by a vector $(\alpha_i, \beta_{i1}, \ldots, \beta_{iK})$. $\epsilon_{ij}$ is the unobserved effect.

Consumers have varying taste for vehicle different attributes. For example, an individual from a high-income family is more likely to be less sensitive to price than one from a low-income
family. Also, an individual who drives more frequently is more likely to favor a vehicle that is more energy efficient. To capture consumers’ varying taste, we assume that the coefficient for vehicle characteristic depends on consumer attributes, including family income and personal demographics:

$$\beta_{ik} = \hat{\beta}_k + \sum_d r_{id} \cdot \beta^d_{ik} + v_{ik} \cdot \beta^u_k$$

where \((r_{i1}, \cdots, r_{iD})\) is a vector consumer \(i\)’s attributes. \(v_{ik}\) is the unobserved consumer attributes.

Each consumer chooses the vehicle which brings him the highest utility. Hence, the set attributes of consumers who will choose the vehicle \(j\) is

$$P_j = \{(\alpha_i, r_{i1}, \cdots, r_{iD}, v_i) | u_{ij} \geq u_{ih}, \forall h \neq j\}.$$  

Then we relate the set \(P_j\) with the distribution of consumer income and demographics from the US census data and the market sales data.

4. Combining with Census Data

One novel design in our study is that we combine the data of auto sales in the market level with census data that describes the consumer attributes. Incorporating the survey data into our study has several advantages:

Firstly, it allows us to reach more precise results on how consumers make their decisions. The theoretical model described above is rendered unrestrictive because the results are based the sales data, which can be rationalized by multiple combinations of consumer attributes and consumer taste. As consumer tastes are not directly observable, adding restrictions on consumer attributes by matching the distribution \((r_{i1}, \cdots, r_{iD})\) to the empirical distribution from the census data increases the plausibility of the results.

Secondly, it provides richer insights as to how consumers from various demographic backgrounds contribute to differ in their decisions. For example, we would be able now how people with different annual income differ in their attitudes toward hybrid cars. While consumers whose incomes are relatively low might be more sensitive to price fluctuation – which means they favor hybrid car, they might not be able to afford one. To understand how they balance such a tradeoff, we need to cluster the underlying population according to the distribution generated from the census data.

Thirdly, it guides the policy implication. Which population is most sensitive to tax credits? Which population is most likely to eschew from purchasing hybrid car? By answering these questions with the census data, we could design more effective policy.
5. Current Progress

6.1 We developed the theoretical model that describes how consumer demand is affected by external policy, such as tax credits. It starts from a generic characterization of consumer’s preference and aggregate the individual choice into the market demand.

6.2 We have finished the data collection, including the auto sales data, census data, and auto characteristics data. These three datasets are now combined with each other.

6.3 We programmed the estimation model of market demand. The program is written under the environment of Stata. It includes the estimation of consumers’ decision process on purchasing vehicles and the aggregation of market data.

7. Remaining Progress

We expect to fully implement the estimation and get the results in the next stage. It includes the estimation of consumers’ demand for auto characteristics and robustness checks which examines how the results are plausible and valid under various assumption and restrictions.

Next, we will conduct counterfactual analysis on how alternative policy and external environments (namely gas price levels) affect the demand for hybrid vehicles. These policies will be evaluated according to different criterions.
Finally, we aim to write a publishable paper that fully describes the project, which includes the scope and background, data collection, model specification, estimation procedure, estimation results and policy implications.