The Household Vehicle Portfolio: Implications for Emissions Abatement Policies

James Archsmith
UC Davis

Kenneth Gillingham
Yale

Christopher Knittel
MIT

David Rapson*
UC Davis

November 16, 2016

PRELIMINARY AND INCOMPLETE.
Please do not cite or circulate.

Abstract

This paper quantifies the extent to which multi-car households exhibit preferences for a diversified vehicle portfolio. We deploy a novel identification strategy to examine how an exogenous change in the fuel economy of a kept vehicle affects a household's choice of a second vehicle purchased. This has potential implications for the relative long-run effectiveness of greenhouse gas abatement initiatives such as fuel economy standards and a price on carbon. Results indicate a strong preference for a diverse portfolio in fuel-economy. They highlight the importance of our instrumental variables approach and the pitfalls that arise when using between-household, rather than within-household, variation. Our results imply that the portfolio effect will exert a strong force that may erode a substantial portion of expected gasoline savings from higher fuel economy standards.

*Archsmith: PhD Candidate in Economics, UC Davis, jesmith@ucdavis.edu. Gillingham: Assistant Professor, Yale University and NBER, kenneth.gillingham@yale.edu. Knittel: William Barton Rogers Professor of Energy Economics, Sloan School of Management, and Director of the Center for Energy and Environmental Policy Research, MIT, and NBER, knittel@mit.edu. Rapson: Associate Professor of Economics, UC Davis, dsrapson@ucdavis.edu.
This paper has benefited from conversations with seminar participants at the Energy Institute @ Haas Summer Energy Camp. We thank James Archsmith for excellent research assistance. We gratefully acknowledge financial support from the California Air Resources Board. The statements and conclusions presented are those of the authors and not necessarily those of the California Air Resources Board. The mention of commercial products, their source, or their use in connection with material reported herein is not to be construed as actual or implied endorsement of such products.

1 Introduction

Understanding the demand patterns for vehicles is important for a number of policy issues, most notably fuel economy standards or gasoline taxes on vehicle choice. Empirical models used to analyze the costs and benefit of such policies have typically assumed away the multi-vehicle nature of many households. That is, these models assume that consumers choose only only vehicle; or, alternatively, that the choice of each vehicle in a household is independent of the others.

There are likely to be two sources of interdependence. The first is that households vary in their "type." This source of dependence is *implicitly* captured in empirical models that allow for variation in the willingness to pay for vehicle attributes. For example, the choice of fuel economy across vehicles within a household will be positively correlated for a household that particularly values horsepower. An alternative source of interdependence exists if households have a taste for diversification. In this case, a household endowed with a high horsepower vehicle will favor a more fuel-efficient second vehicle.

The presence of this second form of interdependence can alter the predictions from policy counterfactuals. For example, suppose a policy were to increase the chosen fuel economy of vehicles for a given household at time t. Counterfactual predictions based on models that assume independence would generate a naive measure the fuel economy benefits. If households prefer diversification, the actual fuel economy benefit from the policy across the entire household portfolio would be smaller, as the fuel economy of the next vehicle purchased would fall.

In this paper, we use panel data on the portfolio of vehicles within a household to estimate how a household's choice of vehicle depends on the other vehicles owned by the household. In particular, our data track the portfolio of vehicles, for a particular household, over time. Our empirical strategy estimates how the fuel economy of a newly-added vehicle depends on the fuel economy of the vehicle already held by the household. Our identification strategy relies both on the richness of the panel, which allows us to control for household-level fixed effects, as well as a novel instrumental variables approach to control for the endogeneity of the fuel economy of the existing vehicle.

The ideal experiment for our research question would randomly assign the "kept" vehicle to households in the market for a new or used vehicle and then observe the relationship between the fuel economy of this kept vehicle and the fuel economy of the newly-acquired vehicle. Since this ideal experiment is obviously not possible, our identification strategy must overcome two sources of endogeneity stemming from the non-random assignment of the kept vehicle. The first is the choice of which vehicle to replace. Since the household preference for particular features of a multi-car portfolio will directly inform the decision of which car to keep or drop, there is an identification challenge in estimating the portfolio itself using observational data. The second is related to the presence of unobserved household preferences. Household fixed effects can address time invariant unobserved preferences, but there would still be a concern if these preferences change over time.

We expect that these time-varying preferences would generate a positive correlation between the fuel economy of the kept and newly-acquired vehicles.

We employ two sets of instruments to account for these potential sources of bias. The first set of instruments are derived from the observation that changes in the relative price of cars in a portfolio systematically affect the probability that the lowest fuel economy car is dropped. We discuss and present three instruments that rely on this feature of the choice setting, with our preferred instrument being deviations in expectation of the change in relative vehicle prices at the time when the kept car was initially purchased. The second instrument is the gasoline price at the time of the purchase of the kept vehicle. A number of papers (Klier and Linn (2010), Busse et al. (2013), Gillingham (2011)) have shown that vehicle purchase behavior is influenced by contemporaneous gasoline prices. Given the results of this literature, we would expect that the fuel economy of the kept vehicle is influenced by the gasoline price at the time of that purchase—something we confirm in our own data. We argue that the instrument satisfies the exclusion restriction because after controlling for current gasoline prices, we would not expect past gasoline prices to influence the choice of the new vehicle. This, we grant, rests on the assumption that consumers are not using this past gasoline price in the formation of the expected gasoline prices and we have adequately controlled for serial correlation in the residuals.

We find evidence that households value diversification. The greater is the fuel economy of the kept car, the lower the fuel economy of the purchased car. We show this using both a continuous measure of fuel economy, as well as by estimating the probability a household purchases a vehicle in the upper and lower quartiles of the fuel economy distribution. Increases in the fuel economy of the kept car reduces the probability the household purchases a car in the lower quartile of gallons per mile, while such increases reduce the probability the household buys a car in the upper quartile. Changes in gasoline prices affect the preferences for diversification in intuitive ways. As gasoline prices increase, the effect of the fuel consumption of kept vehicle and the probability of buying a car in the lower quartile of fuel consumption becomes even more positive. In contrast, as gasoline prices increase, the effect of the fuel consumption of kept vehicle and the probability of buying a car in the upper quartile of fuel consumption becomes even more negative.

To gauge the importance of the portfolio effect, we use our results to estimate the net effect of an exogenous increase in the fuel economy of the kept vehicle. We calculate the decrease in the fuel economy of the newly purchased vehicle when we increase the fuel economy of the kept vehicle by 10, 25, and 50 percent, across gasoline prices of $2.00, $3.00, and $4.00. These calculations suggest that the portfolio effect can have large consequences of the net effect a one-time increase in fuel economy; between 75 to 95 percent of the fuel savings from increasing the fuel economy of the kept vehicle are eroded from the resulting decrease in fuel economy of the newly purchased vehicle.

The remainder of the study proceeds as follows. The next section describes the household vehicle choice problem and outlining a simple theoretical model (Section 2). We then describe our datasets, the restrictions that determine the sample used for our empirical tests, our identification

strategy and empirical approach (Section 3). We then present our results and their economic importance (Section 4). We conclude with a brief discussion of the implications for policymakers and empiricists (Section 5).

2 Context and Model

We begin by developing a simple economic framework to fix ideas and motivate our empirical work. Consider a household who has a vehicle and is considering adding a second vehicle. They may have just sold their second car, or they may be adding a new car to the household's vehicle portfolio. For simplicity, ignore the outside option of not purchasing a second vehicle. Consider a standard discrete choice framework with a random utility model. The household is the decision-maker in this framework; we abstract from issues of within-household bargaining.

Let the characteristics of a vehicle be given by the vector θ_V, where $V \in \{A, B, ...\}$ is the vehicle type. Vehicle types may be defined broadly, such as the class of vehicle (e.g., SUV or small car), or at a finer level, such as at the make-model level. Suppose the household has a vehicle of type A to start.

The household receives utility based on the characteristics of each type of vehicle, and may also receive utility from having a diversity of vehicles, which allows them to optimize their use of the vehicles (e.g., use the larger one for hauling goods and the more efficient one for the longer commutes). Let the contribution to utility from the diversity of the portfolio be given by Γ_{V_1,V_2}, where V_1 and V_2 are the vehicles types for the first and second car respectively.

The indirect utility for household i starting with a vehicle of type A and purchasing a vehicle of type B is thus given as:

$$u_i^{AB} = f(\theta_A) + f(\theta_B) + \Gamma_{AB} - \alpha(p_{A1} + p_{B2}) + \epsilon_i,$$

where $f(.)$ is a function that converts characteristics into consumer utility, and p_{Vj} is the remaining "present value lifetime ownership cost" for a vehicle of type V and order in the household j, where $j = 1$ refers to the vehicle the household already holds and $j = 2$ refers to the new vehicle. To understand the present value lifetime ownership cost, note that for p_{B2} the ownership cost includes the purchase price in addition to the future costs of fuel and maintenance, while for p_{A1} the ownership cost is just the future fuel and maintenance costs. These can be thought of as expectations based on the future expected driving of each of the vehicles, which need not necessarily be the same across vehicle types. α is the marginal utility of money.

In making the choice of which vehicle type to buy for the second vehicle, the household will compare u_i^{AB} to the utility from all other options. For example, suppose the household with vehicle of type A already is considering the option of buying another vehicle of type A. The utility u_i^{AA} would then be given by:

$$u_i^{AA} = f(\theta_A) + f(\theta_A) + \Gamma_{AA} - \alpha(p_{A1} + p_{A2}) + \varepsilon_i.$$

This equation allows for Γ_{AA}, although this might plausibly be assumed to be 0, for there may be no added utility from having a diverse portfolio if a household buys two of the same type of vehicle. One story for why $\Gamma_{AA} > 0$ would be that by having two of the same type of vehicle, the household receives additional utility by showing peers that they identify with a particular type (e.g., two hybrids showing the household is eco-friendly or two pickups showing the household is "tough").

2.1 Implications for vehicle choice

This paper is about how consumers choose their portfolio of vehicles. In other words, in this simple setting, it is about how the kept vehicle (i.e., vehicle 1 in this setting) influences the choice of the second vehicle.

Continuing the thought experiment where the consumer can only choose options A and B for the second vehicle, we can consider the conditions under which A or B is chosen. Specifically, the household chooses portfolio AB rather than AA if $u_i^{AB} > u_i^{AA}$, which is equivalent to (assuming $\Gamma_{AA} = 0$)

$$f(\theta_B) + \Gamma_{AB} - \alpha p_{B2} > f(\theta_A) - \alpha p_{A2}. \tag{2.1}$$

This simple inequality indicates that the household will choose B when the net utility of the vehicle characteristics, lifetime ownership cost, and portfolio effect from B dominate the net utility of the vehicle characteristics and lifetime ownership cost from A. Rewritten differently, we have

$$\Gamma_{AB} > f(\theta_A) - f(\theta_B) + \alpha(p_{B2} - p_{A2})$$

This states that if the added utility from having a portfolio is greater than the difference in utility from the characteristics of the two types of vehicles plus the difference in the lifetime ownership cost of purchasing the vehicles of the two types (converted to utility terms), then the household will choose a vehicle of type B. In other words, B will be chosen if a positive effect from portfolio diversity is larger than the household "type" effect due to the household valuing the characteristics and lifetime ownership cost. This can be rephrased as an empirically testable prediction:

The portfolio effect will dominate if we observe the household choosing vehicle B when the first vehicle is A (as in the setting described so far), but the "type" effect will dominate if we observe the household choosing vehicle A when the first vehicle is also A.

This simple model provides further insight on the decision process. Let $g(.)$ be the distribution of utilities in the population. Then the simple choice between A and B for the new vehicle, the choice probabilities are given as follows:

$$Pr_{AB} = \int_{\mathbf{u}} I(u_{AB} > 0) I(u_{AB} > u_{AA}) dg(\mathbf{u}),$$
$$Pr_{AA} = \int_{\mathbf{u}} I(u_{AA} > 0) I(u_{AA} > u_{AB}) dg(\mathbf{u}).$$

2.2 Changes in choice with gasoline price or fuel economy standards

This simple model also lends itself to a set of policy-relevant comparative statics.

Changes in Gasoline Prices

Consider a permanent increase in gasoline prices, with perfect foresight of this change. This will imply that p_{B2} and p_{A2} will change based on the relative fuel economy of the two vehicle types. Such a change in the market will have two effects:

1. Direct Effect: The probability of the choosing the higher fuel economy vehicle will increase.
2. Indirect Effect: The relative prices of new vehicles in equilibrium will change, so that higher fuel economy vehicles will increase in price relative to others.

The indirect effect will work in the opposite direction as the direct effect, and depending on the marginal utility of income α, may even dominate. We will see the net of these two effects in the empirics.

Changes in Fuel Economy Standards

Consider the implementation of fuel economy standards. These will have a short-run and long-run effect. The long-run effect is through changing the direction of innovative activity by automobile manufacturers. The short-run effect would imply changing the relative prices of vehicles of different fuel economies, consistent with fuel economy standards putting a shadow price on fuel economy. Focusing only on the short-run effect, this would imply that p_{B2} and p_{A2} would change, so that the higher fuel economy vehicle would be relatively less expensive. This would have a similar effect to the "direct effect" from a permanent gasoline price change.

3 Data and Identification

The cornerstone of our dataset is the universe of California vehicle registration records that occurred from 2001-2007.[1] The DMV dataset includes every vehicle registered under the residential designation code. In California every vehicle must be registered annually. Each record includes the registrant's US Census block group identifier, the 17-digit vehicle identification number (VIN) that uniquely identifies the vehicle, that year's registration date, the date when the vehicle was

[1]We thank the California Department of Motor Vehicles (DMV) for making these data available for research.

last sold, and various other information. This information allows us to construct a household-level panel dataset vehicle ownership.

Basic vehicle attributes (e.g. horsepower, weight, etc) are available via a VIN decoder that we purchased from DataOne Software. We augment the decoder to include vehicle fuel economy, which is available from the US Environmental Protection Agency. Vehicle-miles traveled are available for each VIN whenever the vehicle is sold and upon receiving biannual Smog Check certification. We thus have a average measure of miles traveled by each vehicle and, by extension, each household for each year in our sample. The coarseness of these data are not optimal for examining high-frequency effects of VMT-switching between vehicles in response to changes in gasoline prices. Nonetheless, gasoline prices are a variable of interest in this study, since they affect the household's optimal portfolio of vehicle fuel economy. Our gasoline price data are at the county-month level.

3.1 Describing the Sample

In each year households are characterized by the starting and ending number of vehicles in their portfolio. In year t a household's starting portfolio size N^s is the number of vehicles registered in that year. If the household registers exactly N^s vehicles in year $t+1$ or $t+2$ then the ending portfolio size N^e in year t is N^s. If the number of vehicle registered in years $t+1$ and $t+2$ are identical, but not equal to N^s then the ending portfolio size is the number of vehicles registered in the later years.[2]

[Table 1 about here]

Table 1 shows the distribution of household portfolio transitions. Specifically, rows indicate the number of cars in year t, and columns indicate the number of cars in $t+1$. The table represents all possible household transitions. The large mass on diagonals indicates that many households do not increase or decrease the number of cars that they register from year to year. A careful interpretation of "0" is necessary: a household with 0 cars is not in our dataset, so transitions from 0 occur when a Californian household without a car in t registers one in $t+1$, or with observationally-equivalence, a household moves to California from another state. Similarly, transitions to 0 occur either when a household sells all of its registered cars, or if it exits the data via a move to another state or a dissolution of the household.

Many of the regressions that follow are estimated using a sample of two-car households that replace one of their cars, a sample which we call "2x2 replacement households".[3] While other

[2] We examine one and two years in the future as a household that may register more cars in one year than they ever owned simultaneously. For example, consider a household that owns two cars in year t. In year $t+1$ they re-register both previously owned vehicles ans the registrations expire. Then, toward the end of the year, they sell one vehicle and replace it with a new one (which requires registration of the new vehicle). This household has registered three unique vehicles in year $t+1$ but only ever owned two at any given time. In year $t+2$, barring the purchase of yet another new vehicle, the household would return to registering two vehicles.

[3] We define a household as replacing one vehicle if the starting (in year t) and ending (in year $t+1$ or $t+2$) portfolios differ by one vehicle. The household may conduct multiple vehicle transactions, as long as one of the two

transitions are certainly interesting, two-car replacement households provide the cleanest experiment. Households increasing the number of cars in their portfolio are likely to be experiencing an unobserved development that increases their demand for transportation (e.g. having a baby). Furthermore, it is unclear how to characterize the channels through which the portfolio of households with more than two cars affects replacement decisions. Does a portfolio effect for those households operate via the highest-VMT kept car, or the newest? Or must the portfolio effect be defined in a higher dimension? Given that no clear answer exists to these questions, we choose the simple path of focusing on the two-car replacement households.

Table 2 shows summary stats for all 2x2 replacement households, including segmentation based on the fuel economy of the bought car. Households that purchase relatively fuel efficient vehicles (gallons per mile quartile 1) tend to keep relatively fuel efficient cars as well. The converse is true for households buying fuel inefficient vehicles, suggesting that households may have an overall preference for either high- or low-GPM cars.

[Table 2 about here]

Many analyses that follow examine use quartile of fuel economy to describe bought and sold cars. The GPM cutoffs are presented in Table 3, along with their corresponding MPG analogs.

[Table 3 about here]

3.2 Identification

To understand the challenges associated with identifying the portfolio effect, we consider a thought exercise. For a given two-car household that replaces one of its vehicles, we would like to know the effect that randomly dropping one of the cars and exogenously perturbing fuel economy of the "kept" car has on the choice of fuel economy of the "bought" car. That is, we would like to randomly assign one car to be the "kept" car, and to randomly assign it a fuel economy (f^k) to see how changes in f^k affect the household's observed choice of f^b, the GPM of the car purchased. This is what we mean when we refer to the "portfolio effect". There are two identification challenges to operationalizing this thought experiment to retrieve an estimate of the portfolio effect in our observational dataset. We propose instrumental variables to address each.

Identification Challenge 1: Which Vehicle to Keep? As described earlier, our sample isolates two-car households that replace one of their cars with another. In general, the choice of which car to keep is endogenous and many potential stories could be told about preferences and conditions that would lead to one or the other of the cars being kept. Such a choice is inconsistent with our thought experiment of randomly assigning the household its f^k. However, our data offer several appealing instrumental variables.

vehicles appears in both the starting and ending portfolios. We do not consider households where both vehicles in the two-vehicle portfolio change as the relative timing of each purchase becomes important for defining the portfolio at the time of each vehicle's purchase.

A valid instrument will provide exogenous variation in the process that determines which of the household vehicles is kept and which is replaced. The exclusion restriction requires that the instrument affect the household's choice of f^b only indirectly, through the choice of which car to keep. We assert that variation in the price differential between the kept and dropped car contains such identifying variation. There are three functions of the price differential that we use. For exposition, let P_t^k and P_t^d be the average retail value of the kept and dropped cars, respectively, at the time when the car is dropped (t). The first candidate instrument is the price difference at time t: $\Delta P^{kt} = P_t^k - P_t^d$. One might be concerned that attributes of the car that are correlated with both the choice of which car to drop and the price difference, which would violate the exclusion restriction.

The second instrument is the change in price differences between time t and time 0, when the kept car was purchased. That is $\Delta\Delta P^{kd} = (P_t^k - P_t^d) - (P_0^k - P_0^d)$. To the extent that market forces are exogenous to portfolio preferences, this instrument has promise. However, one may be concerned that the change in relative prices was expected by the buyer in time t, and thus potentially correlated with preferences in time t as well.

The third candidate instrument addresses the above concerns by extracting only the portion variation in the price difference-in-difference that deviates from expectations. We assume that households form expectation using lagged 1-, 3-, and 5-year depreciation rates at the make-model-modelyear level, and project these into the future. Deviations from these projections are what we refer to as the "deviation difference-in-difference". We find it difficult to come up with a violation of the exclusion restriction for this instrument. Recall that the concern is that a correlation exists between portfolio preference exhibited in the initial purchase of the kept car and the instrument. Relying on an instrument using market-level changes in relative prices that arise only after the purchase of the kept car would be problematic only if those market level changes were correlated with individual household preferences over the vehicle portfolio.

[Figure 1 about here]

Figures 1 (a)-(c) display the reduced form relationship between these price differentials and the probability of dropping the least expensive car in the initial portfolio. Each of the instruments appears to have power. Since the relationships are non-linear, in our estimation we deploy them as third-order polynomials in their respective first stages.

Identification Challenge 2: Omitted Variables. The household's choice of f^k may be influenced by many factors that are unobservable to the researcher. These may include unobserved car attributes that are correlated with GPM (e.g. safety via weight) or unobserved household attributes (e.g. features of commutes). Of particular interest in our setting is fuel economy, and the confounding effect that unobservables may have on f^k. To address this identification challenge, we follow an instrumental variable approach and control for time-invariant household preferences via household fixed-effects.

Our preferred instrument for addressing omitted variables is the price of gasoline at the time of the kept car purchase, $p^{gas}_{it_k}$. Both theory and evidence (e.g. BKZ) demonstrate that households consider future operating costs of the vehicle in their purchase decision. Changes in California gasoline prices are exogenous with respect to the household choice, vary extensively over the time period of our data, and alter the expected lifecycle cost of vehicles according to each vehicle's fuel efficiency. Based on this logic, when gasoline prices are high at the time of the kept car purchase, we would expect the household to purchase a more fuel efficient car than when gasoline prices are low (as also demonstrated in BKZ). The price of gasoline at the time of the kept car purchase thus provides exogenous variation in the potentially endogenous variable of interest, f^k.

Recall that the relationship between f^k and f^b is theoretically ambiguous. A preference for diversification in the household portfolio will lead to a negative correlation, but complementarity between attributes associated with fuel economy may lead to a positive correlation. By extension, the relationship between $p^{gas}_{it_k}$ and f^b may also appear to be positive or negative.

The reduced form relationship between the instrument and our outcome variable of interest, f^b, is presented in Figures 2a and 2b. Many factors influence a consumer's choice of vehicle attributes, including f^b, so a plot of the raw data reveals little about the underlying relationship between our variables of interest. Instead, we present the variables after partialing out other covariates. The x-axis and y-axis are the residuals retrieved from regressing $p^{gas}_{it_k}$ and f^b, respectively, on covariates.

[Figures 2a and 2b about here]

A clear relationship emerges. For new cars, we observe a *lower* (conditional) f^b at higher levels of (conditional) $p^{gas}_{it_k}$. The relationship is most clear in the Kernel regression line. The relationship is reversed for used cars: a *higher* (conditional) f^b is observed at lower levels of (conditional) $p^{gas}_{it_k}$. For the purposes of our empirical approach, what is important here is that we observe a relationship between the instruments and the outcome variable of interest, which (if you believe the exclusion restriction) implies a strong first-stage.

Further Consideration: Within-Household Variation. We further refine our identification strategy to address potential concerns that first-order household preferences for cars with high (or low) fuel economy may overwhelm our ability to identify the potentially second-order preference for a diversified fuel economy portfolio. The richness of our panel dataset allow us to deploy household fixed effects to control for time-invariant unobservable preferences such as this. The importance of these controls can be seen via a simple example. Suppose that there are two types of households. They both prefer a diverse vehicle portfolio, but one (say household A) has an overall preference for gas guzzlers and the other (household B) for fuel efficiency. Examining each household's portfolio may reveal that household A holds cars that are both in the highest GPM quartile, whereas household B holds cars that are both in the lowest GPM quartile. Were we to randomly remove one of the cars from each portfolio, they would each be left with a car in the GPM quartile of their preference. They would also be likely to purchase a new car that is also in that

GPM quartile. On the surface, it would appear as though the households have a low preference for a diversified vehicle portfolio. However, that may be a false conclusion. Were we to examine GPM *within* the preferred quartile, we may discover that the household prefers diversification within that range. Using household fixed effects as controls allows our empirical approach to reveal the true impact of an exogenous marginal change in the fuel economy of the kept vehicle on the (marginal) fuel economy of the vehicle purchased.

Identifying household fixed effects requires observing at least two transactions per household, which imposes a restriction on our viable sample. Figures 3a - 3b present histograms of the number of transactions per household under various sample restrictions. It reveals that, while many households must be excluded to estimate specifications with household fixed effects, we are still left with approximately 235,000 households in the IV specification that includes household fixed effects.

[Figures 3a to 3b about here]

3.3 Regression Specifications

The basic regression strategies examine the relationship that fuel economy of the kept car has on the chosen fuel economy of the bought car. The dependent variable is thus either fuel economy of the bought car itself (f^b_{it}), or quartile indicators of that variable. Regressors of interest include gasoline price at the time of purchase, fuel economy of kept car (f^k_{it}), and their interaction. In addition, we include a term, $\mathbb{1}\{\Delta f^{kd} > 0\}$, that distinguishes which car was dropped from the initial portfolio, the low- or high-GPM car. Specifically, $\Delta f^{kd} = f^k_{it} - f^d_{it}$ in our main specification, such that $\mathbb{1}\{\Delta f^{kd} > 0\} = 1$ indicates that the car with the lower fuel economy in the initial portfolio is kept.

Most of the regression results that follow are retrieved from estimating a linear model of the probability of purchasing vehicles in a given MPG quartile. For ease of exposition of the results, and to allow our focus to rest on what happens in the top and bottom quartile, we combine vehicles in the 2nd and 3rd quartiles are into a single category, "med". The baseline specification is

$$Pr\left(q(f^b_{it}) = s\right) = \beta_0 + \beta_g p^{gas}_{it} + \beta_f f^k_{it} + \beta_{gf} p^{gas}_{it} \times f^k_{it} + \beta_{df} \mathbb{1}\{\Delta f^{kd} > 0\} + \alpha_X X^k_{it} + \varepsilon_{it} \qquad (3.1)$$

where the dependent variable, $Pr\left(q(f^b_{it}) = s\right)$, equals one if f^b_{it} falls within the range of quartile $s \in \{1, med, 4\}$. We also estimate a continuous model where the dependent variable is f^b_{it}, keeping the rest of the specification as presented in equation 3.1. Fuel economy of the vehicles bought (b) and kept (k) by household i in time t are denoted f^b_{it} and f^k_{it}; i's contemporaneous gas price in t is p^{gas}_{it}, whereas $P^{gas_k}_{it}$ is the price of gasoline *at the time household i purchased the car that it keeps in time t*. Control variables, denoted X_{it}, include vehicle attributes (e.g. class, make, value, age), nonparametric time controls (year and month-of-year fixed effects) and household/demographic (household fixed effects and county-level unemployment).

The IV specifications deploy instruments for the indicator of the kept vehicle's rank in fuel economy within the portfolio ($\mathbb{1}\{\Delta f^{kd} > 0\}$), the kept vehicle fuel economy (f^k_{it}), and the interaction of gas price and fuel economy ($p^{gas}_{it} \times f^k_{it}$). In each specification, we instrument using the gas price at the time the kept vehicle was purchased ($P^{gas_k}_{it_k}$) and that gas price interacted with the current gas price ($p^{gas}_{it} \times P^{gas_k}_{it_k}$). We augment this set of instruments with the instruments based on vehicle price differences that were briefly described in Section 3.2 on identification to estimate the following system of endogenous variables:

$$\mathbf{Z}_{it} = \left[\begin{array}{ccc} f^k_{it} & p^{gas}_{it} \times f^k_{it} & \mathbb{1}\{\Delta f^{kd} > 0\} \end{array} \right]'$$

We now describe the vehicle price difference instruments precisely. In "Price Difference" specification, we include the difference in the current resale value of the kept and sold vehicles ($\Delta P^{kd}_{it} = P^k_{it} - P^d_{it}$) as an additional instrument. The "Price Difference-in-Difference" specification uses the change in value for the kept and dropped vehicles between the point the vehicle was purchased and the current time period: $\Delta\Delta P^{kd}_{it} = (P^k_{it} - P^k_{i0}) - (P^d_{it} - P^d_{i0})$.

The third instrument, which we call "Price Deviation Difference-in-Difference", is constructed from the deviation of the difference between the kept and dropped vehicles relative to their expected deprecation rates at the time of the kept car purchase. For each of the kept and dropped vehicle we estimate the households expectation of annual vehicle depreciation using depreciation of similar vehicles over the previous five years. Specifically, for vehicle make m and model year y, and value $V_{m,y,t}$ in year t, the expected depreciation is[4]

$$\mathbf{E}[Dep_{m,y,t}] = \left(\prod_{s=1}^{5} \frac{V_{m,y-s+1,t-s+1} - V_{m,y-s,t-s}}{V_{m,y-s,t-s}} \right)^{\frac{1}{5}} \tag{3.2}$$

We can then calculate the deviation from this expected depreciation rate for each car in the portfolio, and construct the "Price Deviation Difference-in-Difference" instrument. Assuming vehicle j has resale value $P_{j,t}$ in year t, this is:

$$\Delta\Delta V^{kd}_{it} = (P^k_{it} - \mathbf{E}[Dep^k_{it}] \cdot P^k_{i,t-1}) - (P^d_{it} - \mathbf{E}[Dep^d_{it}] \cdot P^d_{i,t-1}) \tag{3.3}$$

The set of three price difference instruments is $W = \{\Delta P^{kd}_{it}, \Delta\Delta P^{kd}_{it}, \Delta\Delta V^{kd}_{it}\}$. The first stage

[4] As a more concrete example, for a household in year $t = 2005$ owning a 2002 Honda Civic, the expected depreciation is the geometric mean annual depreciation rate of 2001 Hondas in 2004, 2000 Hondas in 2003, 1999 Hondas in 2002, etc.

thus consists of the following system of three equations for each of the instruments $w \in W$:

$$\begin{aligned}\mathbf{Z}_{it}^{w} = \mathbf{\Gamma}_0 + \mathbf{\Gamma}_g p_{it_k}^{gas} + \theta_P P_{it}^{gas_k} + \mathbf{\Gamma}_{gg} p_{it}^{gas} \times p_{it_k}^{gas} \\ + \sum_{dc \in CLASS} \mathbf{\Gamma}_{dc} \mathbf{1}[CLASS_{it}^{dropped} = sc] \\ + \mathbf{\Gamma}_w w + \mathbf{\Theta} \mathbf{X}_{it} + \mathbf{\Xi}_{it}\end{aligned} \tag{3.4}$$

4 Results

The objective of this section is to present and justify our empirical approach, illuminate the effect of key variables on the choice of bought car fuel economy, and present some simple counterfactual analyses that demonstrate the policy-relevance of our findings. The section is comprised of two main parts.

First, we show results from various regression specifications. This allows us to demonstrate the importance of our instrumental variables approach and the inclusion of household fixed effects, both of which qualitatively alter key coefficient estimates. We then present marginal effects of gas prices and kept car fuel economy on bought car fuel economy, which reveals interesting features of household preferences for a diversified portfolio.

In the second subsection, we present results from counterfactual analyses. Specifically, we exogenously perturb the fuel economy of the kept vehicle, which is roughly what a successful fuel economy standard would do.[5] We observe the extent to which potential gasoline consumption reductions are either magnified or eroded due to portfolio considerations will become apparent.

4.1 Regressions and Marginal Effects

Tables 4 and 5 present the baseline regression results from new and used car purchases, respectively. The left-most column in each tables corresponds to estimates using the Price Difference IV from Equation **??**. Column 2 presented estimates using the Price DiD IV shown in Equation **??**. Finally, Column 3 is our preferred specification, deploying the Price Deviation DiD IV from Equation 3.4. The four panels correspond to the continuous dependent variable regression (Panel A), the linear probability model on the highest fuel economy quartile purchases (Panel B), the linear probability model on second and third fuel economy quartile purchases combined (Panel C), and the linear probability model on purchases of cars in the lowest fuel economy quartile (Panel D).

[Tables 4 and 5 about here]

Looking across the columns of Table 4 and 5 the specifications are robust to the specific choice of instruments. First stage power decreases from left to right as the instruments discard more

[5] A fuel economy standard will cause new car buyers to purchase more fuel efficient cars, on average. The question of interest in this counterfactual is whether and to what extent this effects the desired choice of fuel economy in the next vehicle purchase.

information contained in the relative resale prices of the kept and dropped vehicles. However, even our preferred instrument results in a powerful first stage; the Cragg-Donald minimum eigenvalue statistic from the system of first-stage IV regressions is reported below the table, and it reflects a very strong first stage for both new and used cars.

The inclusion of Δf^{kd} is important, reflecting the difference between dropping the lower- or higher-GPM vehicle in the portfolio. In Panel B, for example, a negative coefficient on Δf^{kd} would indicate that, for households dropping their fuel-efficient (low-GPM) car, a larger the difference between GPM in their initial portfolio leads the household to buy a low-GPM replacement with more likelihood (i.e. the household replaces the dropped car with a car of like kind). In Panel D, the logic is identical, but the sign is reversed: a *positive* coefficient on Δf^{kd} implies the fuel economy of the replacement car is higher than the kept car if the dropped car fuel economy was also higher.

The coefficients on Δf^{kd} are thus important litmus tests for the LPM results. If we were to find that households encouraged to drop their Prius due to unexpected change the relative value of vehicles in their portfolio generally then replaced it with a Hummer, that would be suspicious. This simple example provides a thought exercise for understanding the important role of household fixed effects. When the regression is identified using across-household variation, the evidence indicates that households will tend towards replacing their dropped car with one that is qualitatively similar in GPM to the kept car. In this way, there appear to be some households that prefer fuel efficiency and others that prefer gas guzzlers (presumably due to power, comfort, safety, etc). It is only when we look within the household that the portfolio effect of interest is seen. Despite the possibility that some households prefer low- or high-GPM cars in general, there appears to be a preference for diversity in GPM within that band.

To more clearly understand the effect of key covariates, we now present and discuss their marginal effects (Tables 6 to 7). First we discuss marginal effects of gasoline prices on the fuel economy of new and used cars, followed by the marginal effects of kept car GPM on bought car GPM. For readers preferring a graphical representation, these are shown in Figures 4 to 7.

[Tables 6 and 7 about here]

[Figures 4 - 7 about here]

The effect of gasoline prices on bought car GPM is quite heterogeneous across both new and used cars, and across kept car GPM quartiles. New car GPM is relatively unaffected by changes in gasoline prices, although there is an decrease in the probability of buying a car in the highest-GPM quartile (fuel inefficient) as gas prices rise. Surprisingly, the probability of buying a new car in the top quartile of fuel efficiency does not increase with gasoline prices. There are a multitude of effects on the probability of used car purchases across the distribution, however. This should not be surprising, given the equilibrium effects that gas prices induce on both the the supply- and demand-sides of the market.

Note that the continuous regressions (Panel A) shroud heterogeneous marginal effects that occur across the distribution. There does appear to be a little desire for a diversified portfolio when examining the marginal gas price effect on high- and low-GPM quartile used car purchases. Households with fuel efficient kept cars are more likely to purchase another high fuel-economy car as gas prices rise; similarly, households that keep a car in the lowest fuel economy quartile are less likely to buy a fuel efficient car as gas prices rise. Demand for middle-quartile GPM cars responds in an opposite way. Owners of fuel efficient cars are less likely to buy middle-quartile GPM cars as gas prices rise, whereas owners of low fuel economy cars are more likely to buy a middle-quartile GPM car. The results for low fuel-economy used purchases reflect the general equilibrium effects mentioned above. As gas prices rise, households that keep a fuel efficient car are more likely to buy a gas guzzler, as result that indicates supply-side effects in the used market outweigh demand effects. On the other hand, households that keep a fuel-inefficient car are less likely to buy another one as gas prices rise.

Marginal effects of kept car GPM provide a more direct representation of the portfolio effect. We present marginal effects of kept car fuel economy for new and used cars (and pooled) at different gasoline prices in Table 7. Note that both the level of the coefficients within each bought car GPM quartile, and the gradient of these coefficients with respect to the gas price, are important. Positive coefficients reflect an increasing probability of buying in a given quartile as kept car GPM increases (becomes less fuel efficient).

The overall story is clear: households exhibit a preference for GPM diversification in their portfolio, and that preference increases as the gas price rises. This can be seen clearly in the overall results presented in panel A, where negative coefficients imply an increased demand for buying a fuel efficient car as kept-car fuel efficiency decreases. A similar narrative holds when examining the high- and low-GPM bought car quartiles. Households buy new cars in the lowest GPM quartile (high fuel economy) with a higher probability as their kept car fuel economy decreases. Furthermore, this preference increases as the gas price rises. On the other hand, households have a lower probability of buying fuel inefficient new cars as kept car GPM increases, and this preference also increases as the gas price rises.

The situation for used cars is qualitatively similar. The negative coefficient on the marginal effects on 1st GPM quartile implies that the probability of buying a used, highly efficient car decreases in kept-car GPM at this gas price. We presume that this reflects a supply-side effect: households with fuel efficient used cars are less likely to sell them when gas prices are high, thus shifting upwards along the demand curve for this type of car.

4.2 Counterfactuals (Preliminary)

The results presented above provide a platform for examining how preferences over the household portfolio may affect energy conservation resulting from popular environmental policies. In particular, we view these results as particularly informative about long-run impacts of fuel econ-

omy standards, such as Corporate Average Fuel Economy (CAFE) standards in the United States. CAFE affects the suite of cars available for purchase, and their purchase prices. Once these cars become part of a household's vehicle portfolio, the change in attributes (relative to no CAFE standard) may influence the subsequent choice of vehicle purchased. In particular, if households exhibit a preference for diversification in their portfolio, increasing the fuel efficiency of their kept car will lead to a less fuel efficient second-car purchase.

The counterfactuals that we describe here examine the net effect of an exogenous decrease of 10% and 25% in the kept car GPM (i.e. a fuel efficiency increase) on predicted gasoline consumption. The net effect includes changes in gasoline consumption relating to use of the (now more efficient) kept car, but also changes in gasoline consumption relating to the use of the bought car whose GPM is influenced via the portfolio effect that we estimate above. Throughout the exercise, we assign cars the vehicle-miles traveled (VMT) that are observed in our dataset. We do not adjust these to account for a rebound effect, although in reality, one may exist.[6] Notice that this implies the only changes in gasoline consumption that we estimate will occur due to changes in the (extensive margin) choice of bought car fuel economy that are induced via the portfolio effect. It may well be the case that there is a first-order effect of gas prices on bought car GPM, but we seek to isolate only the portfolio effect in this exercise.

[Tables 8 and 9 about here]

Turning to Table 8, first note the observed gas consumption of the kept and bought vehicles. Each is estimated to consume approximately 550 gallons of gasoline per year (the product of VMT and vehicle GPM). Changes in kept car GPM do have a large effect; a 10% decrease in kept car GPM mathematically reduces kept car gas consumption by 10%, as can be seen in the top rows at each gas price level. The increases in bought car gas consumption reflect the portfolio effect. It is immediately apparent that these effects are quite large, and that the majority of gasoline conservation enjoyed by a kept car GPM improvement are eroded by the household's response via the choice of lower bought car fuel economy. Indeed, our estimates predict that the portfolio effect offsets about two thirds of the fuel savings from increasing the kept vehicle's fuel economy.

The story is qualitatively similar when the bought car is used instead of new. Table 9 reveals that households purchasing used cars have, on average, lower gasoline consumption associated with the vehichle added to the portfolio, but slightly higher gasoline consumption with the kept vehicle. The exogenous GPM changes associated with the kept car are 67% offset by the portfolio effect.

These results are quite startling and, if true, have unfortunate implications for the effectiveness of fuel economy standards as a way to reduce gasoline consumption. While it is clear that these portfolio effects cannot be directly applied to a fuel economy standard, the magnitude of the portfolio response implies that strong forces will be at work, particularly in the used car market,

[6] As fuel economy changes it, in turn, alters the cost per mile traveled. Consumers faced with this change in relative prices may choose a different VMT. See Borenstein (2015) and Gillingham et al. (2016) for more on the rebound effect.

which is not covered by CAFE. Increased demand for used, fuel inefficient cars will occur as a result of increased efficiency from CAFE. The increase in demand will lead used gas guzzlers to be more valuable, and thus more slow to be retired from the fleet (similar to the effect documented in Jacobsen and van Benthem (2015).

5 Conclusions

The effects of a number of polices applied the light duty vehicle market depend crucially on vehicle choice patterns. Typically empirical estimates of vehicle choice assume that the vehicle choices within a household are made independently. We provide show evidence that this assumption does not hold.

Using panel data on the portfolio of vehicles within a household, and a novel instrumental variables approach, we find evidence that households value diversification; exogenous increases in the fuel economy of the kept car lowers the fuel economy of the purchased car. We show this using both a continuous measure of fuel economy, as well as by estimating the probability a household purchases a vehicle in the upper and lower quartiles of the fuel economy distribution. Increases in the fuel economy of the kept car reduces the probability the household purchases a car in the lower quartile of gallons per mile, while such increases reduce the probability the household buys a car in the upper quartile.

We also find that gasoline prices affect the preferences for diversification in intuitive ways. As gasoline prices increase, the effect of the fuel consumption of kept vehicle and the probability of buying a car in the lower quartile of fuel consumption becomes even more positive. In contrast, as gasoline prices increase, the effect of the fuel consumption of kept vehicle and the probability of buying a car in the upper quartile of fuel consumption becomes even more negative.

To understand the economic important of this taste for diversification, we use our results to estimate the net effect of an exogenous increase in the fuel economy of the kept vehicle. These calculations suggest that the portfolio effect can have large consequences of the net affect a one-time increase in fuel economy; between 75 to 95 percent of the fuel savings from increasing the fuel economy of the kept vehicle are eroded from the resulting decrease in fuel economy of the newly purchased vehicle.

References

Borenstein, S. (2015). A microeconomic framework for evaluating energy efficiency rebound and some implications. *Energy Journal*, 36:1–21.

Busse, M. R., Knittel, C. R., and Zettelmeyer, F. (2013). Are consumers myopic? evidence from new and used car purchases. *American Economic Review*, 103(1):220–256.

Gillingham, K. (2011). How do consumers respond to gasoline price shocks? heterogeneity in vehicle choice and driving behavior. Manuscript: Yale University.

Gillingham, K., Rapson, D., and Wagner, G. (2016). The rebound effect and energy efficiency policy. *Review of Environmental Economics and Policy*, 10:68–88.

Jacobsen, M. R. and van Benthem, A. A. (2015). Vehicle scrappage and gasoline policy. *American Economic Review*, 105(3):1312–1338.

Klier, T. and Linn, J. (2010). New vehicle characteristics and the cost of the corporate average fuel economy standard. Resources for the Future Discussion Paper 10-50.

Table 1: Number of Unique Households by Portfolio Size

Start Portfolio Size	End Portfolio Size 1	2	3	4+
1	7,262,111	1,360,594	187,558	75,150
2	1,172,278	4,632,425	839,546	259,098
3	168,745	849,703	2,169,948	675,040
4+	35,810	141,618	381,226	1,489,926

Each cell represents the count of unique households from 2001 to 2007 observed to have the starting portfolio size shown in each row and the ending portfolio size shown in the column. These counts provide a measure of the number of households providing identifying variation in each portfolio cell. A single household may appear in multiple cells if their portfolio changes over time but is counted at most once in each cell. For example, two-car household that replaces one car every year would add one to the count of the (2,2) cell. If instead, that household adds a third vehicle in 2004 and returns to a two-car portfolio in 2006 it would add one to the count of the (2,2) cell, one to the count of the (2,3) cell, one to the (3,3) cell, and one to the count of the (3,2) cell. Each household may have zero, one, or multiple vehicle transactions during this time period.

Table 2: Summary Statistics - 2x2 Replacement Households

	All Households	Bought GPM Qtile 1	Bought GPM Qtile 2 or 3	Bought GPM Qtile 4
N Households	1,452,896	392,168	768,517	413,367
N Transactions	2,004,312	491,010	1,003,044	510,258
Mean Transactions/HH	1.34	1.24	1.24	1.23
Std. dev	(0.676)	(0.504)	(0.612)	(0.491)
Median Transactions/HH	1.00	1.00	1.00	1.00
IQR	[2.00 - 1.00]	[1.00 - 1.00]	[1.00 - 1.00]	[1.00 - 1.00]
Median Bought GPM	0.05	0.04	0.05	0.07
IQR	[0.04 - 0.06]	[0.04 - 0.04]	[0.05 - 0.05]	[0.06 - 0.07]
Median Kept GPM	0.05	0.05	0.05	0.05
IQR	[0.05 - 0.06]	[0.04 - 0.06]	[0.05 - 0.06]	[0.05 - 0.06]
Median Bought MPG	20.23	26.01	19.97	15.19
IQR	[17.20 - 22.60]	[23.80 - 27.80]	[18.80 - 21.20]	[14.20 - 16.20]
Median Kept MPG	19.97	20.59	19.88	19.55
IQR	[17.20 - 22.00]	[17.80 - 23.20]	[17.30 - 21.80]	[16.80 - 21.50]

Table 3: Distribution of observed fuel economy

Percentile	Gallons per Mile (GPM)	Miles Per Gallon (MPG)
25th Percetile	0.045	22.0
Median	0.052	19.3
75th Percentile	0.059	17.0

Table 4: Regression Estimates - New Vehicle Purchases

(a) Panel A: Bought Vehicle Continuous GPM Measure

	Price Diff	Price DiD	Price Deviation DiD
	(1)	(2)	(3)
Gas Price ($/ gal)	0.052	0.031	0.032
	(0.009)**	(0.009)**	(0.010)**
Kept GPM	1.525	0.731	0.349
	(0.398)**	(0.413)	(0.426)
Gas Price × Kept GPM	-0.991	-0.584	-0.596
	(0.177)**	(0.170)**	(0.191)**
Δ GPM > 0	-0.001	0.001	0.000
	(0.001)	(0.001)	(0.001)

(b) Panel B: Bought Vehicle 1st GPM Quartile

	(1)	(2)	(3)
Gas Price ($/ gal)	-0.731	-0.228	-0.977
	(0.342)*	(0.322)	(0.447)*
Kept GPM	-3.671	15.711	-22.396
	(15.081)	(14.315)	(18.676)
Gas Price × Kept GPM	13.608	3.929	17.943
	(6.505)*	(6.092)	(8.498)*
Δ GPM > 0	-0.086	-0.119	-0.055
	(0.033)**	(0.031)**	(0.045)

(c) Panel C: Bought Vehicle 2nd/3rd GPM Quartile

	(1)	(2)	(3)
Gas Price ($/ gal)	-1.110	-1.570	-0.290
	(0.546)*	(0.585)**	(0.547)
Kept GPM	-63.774	-85.723	-8.027
	(22.248)**	(27.381)**	(23.113)
Gas Price × Kept GPM	21.483	30.147	6.316
	(10.398)*	(11.107)**	(10.386)
Δ GPM > 0	0.129	0.156	0.063
	(0.046)**	(0.050)**	(0.055)

(d) Panel D: Bought Vehicle 4th GPM Quartile

	(1)	(2)	(3)
Gas Price ($/ gal)	1.841	1.798	1.267
	(0.488)**	(0.518)**	(0.457)**
Kept GPM	67.445	70.011	30.424
	(19.470)**	(24.383)**	(19.583)
Gas Price × Kept GPM	-35.091	-34.076	-24.259
	(9.290)**	(9.832)**	(8.677)**
Δ GPM > 0	-0.042	-0.037	-0.008
	(0.040)	(0.043)	(0.043)
N	440,809	429,369	348,368
Cragg-Donald stat	145.88	141.44	91.34
Household FE	Yes	Yes	Yes
IV for Kept Vehicle	Base+$\Delta PriceDiD^3$	Base+$PriceDiD^3$	Base+ValueDiD
Subsample	New	New	New

Standard errors robust to heteroskedasticity shown in parentheses.

Table 5: Regression Estimates - Used Vehicle Purchases

(a) Panel A: Bought Vehicle Continuous GPM Measure

	Price Diff (1)	Price DiD (2)	Price Deviation DiD (3)
Gas Price ($/ gal)	0.045	0.027	0.057
	(0.018)*	(0.013)*	(0.022)**
Kept GPM	1.086	0.234	1.532
	(0.806)	(0.504)	(0.909)
Gas Price × Kept GPM	-0.874	-0.523	-1.090
	(0.349)*	(0.245)*	(0.421)**
Δ GPM > 0	0.008	0.008	0.007
	(0.001)**	(0.001)**	(0.001)**

(b) Panel B: Bought Vehicle 1st GPM Quartile

	(1)	(2)	(3)
Gas Price ($/ gal)	4.200	3.076	3.081
	(1.137)**	(0.772)**	(1.117)**
Kept GPM	207.785	118.514	109.659
	(50.188)**	(30.916)**	(46.832)*
Gas Price × Kept GPM	-80.034	-58.007	-58.326
	(21.792)**	(14.729)**	(21.382)**
Δ GPM > 0	-0.546	-0.484	-0.489
	(0.061)**	(0.046)**	(0.065)**

(c) Panel C: Bought Vehicle 2nd/3rd GPM Quartile

	(1)	(2)	(3)
Gas Price ($/ gal)	-11.152	-7.979	-9.952
	(2.535)**	(1.605)**	(2.671)**
Kept GPM	-485.343	-251.088	-332.787
	(111.819)**	(65.004)**	(111.180)**
Gas Price × Kept GPM	213.213	151.329	189.695
	(48.602)**	(30.613)**	(51.132)**
Δ GPM > 0	0.582	0.432	0.544
	(0.134)**	(0.094)**	(0.151)**

(d) Panel D: Bought Vehicle 4th GPM Quartile

	(1)	(2)	(3)
Gas Price ($/ gal)	6.952	4.902	6.871
	(1.618)**	(1.033)**	(1.836)**
Kept GPM	277.558	132.574	223.128
	(71.442)**	(41.549)**	(76.125)**
Gas Price × Kept GPM	-133.179	-93.323	-131.369
	(31.022)**	(19.698)**	(35.146)**
Δ GPM > 0	-0.035	0.052	-0.055
	(0.085)	(0.060)	(0.103)
N	500,882	461,425	364,909
Cragg-Donald stat	39.99	42.48	38.93
Household FE	Yes	Yes	Yes
IV for Kept Vehicle	Base+$\Delta PriceDiD^3$	Base+$PriceDiD^3$	Base+ValueDiD
Subsample	Used	Used	Used

Standard errors robust to heteroskedasticity shown in parentheses.

Table 6: Marginal Effect of Gasoline Price on Bought Vehicle GPM - Preferred Specification

(a) Panel A: Bought Vehicle Continuous GPM Measure

Kept Vehicle GPM	All (1)	New (2)	Used (3)
25th Pctile GPM (0.045)	0.00410	0.00453	0.00717
	(0.00111)**	(0.00147)**	(0.00290)*
Median GPM (0.052)	0.00041	0.00070	0.00015
	(0.00031)	(0.00058)	(0.00054)
75th Pctile GPM (0.058)	-0.00317	-0.00302	-0.00666
	(0.00090)**	(0.00121)*	(0.00254)**

(b) Panel B: Bought Vehicle 1st GPM Quartile

	(1)	(2)	(3)
25th Pctile GPM (0.045)	0.15223	-0.16149	0.42954
	(0.05208)**	(0.06575)*	(0.14765)**
Median GPM (0.052)	0.02598	-0.04597	0.05401
	(0.01608)	(0.02728)	(0.03124)
75th Pctile GPM (0.058)	-0.09651	0.06611	-0.31032
	(0.04205)*	(0.05472)	(0.13038)*

(c) Panel C: Bought Vehicle 2nd/3rd GPM Quartile

	(1)	(2)	(3)
25th Pctile GPM (0.045)	-0.62143	-0.00311	-1.32956
	(0.10266)**	(0.08032)	(0.35277)**
Median GPM (0.052)	-0.07545	0.03756	-0.10822
	(0.03161)*	(0.03294)	(0.07349)
75th Pctile GPM (0.058)	0.45426	0.07702	1.07671
	(0.08235)**	(0.06653)	(0.31155)**

(d) Panel D: Bought Vehicle 4th GPM Quartile

	(1)	(2)	(3)
25th Pctile GPM (0.045)	0.46920	0.16460	0.90003
	(0.07463)**	(0.06636)*	(0.24244)**
Median GPM (0.052)	0.04947	0.00841	0.05421
	(0.02210)*	(0.02574)	(0.04896)
75th Pctile GPM (0.058)	-0.35774	-0.14313	-0.76639
	(0.06005)**	(0.05475)**	(0.21349)**

Marginal effect of the current gasoline price on the probability a household purchases a vehicle in the GPM quartile specified in table section header. Delta method standard errors robust to heteroskedasticity shown in parentheses. *,**,*** denote results significant at the 10%, 5%, and 1% levels, respectively.

Table 7: Marginal Effect of Kept Vehicle MPG on Bought Vehicle GPM - Preferred Specification

(a) Panel A: Bought Vehicle Continuous GPM Measure

Current Gas Price	All (1)	New (2)	Used (3)
Gas Price = 2.00/gal	-0.819	-0.843	-0.648
	(0.140)**	(0.202)**	(0.242)**
Gas Price = 3.00/gal	-1.393	-1.438	-1.738
	(0.235)**	(0.283)**	(0.453)**
Gas Price = 4.00/gal	-1.966	-2.034	-2.828
	(0.369)**	(0.438)**	(0.840)**

(b) Panel B: Bought Vehicle 1st GPM Quartile

	(1)	(2)	(3)
Gas Price = 2.00/gal	-12.832	13.490	-6.992
	(6.902)	(8.860)	(13.288)
Gas Price = 3.00/gal	-32.442	31.433	-65.318
	(11.113)**	(12.649)*	(23.234)**
Gas Price = 4.00/gal	-52.051	49.376	-123.643
	(17.249)**	(19.646)*	(42.632)**

(c) Panel C: Bought Vehicle 2nd/3rd GPM Quartile

	(1)	(2)	(3)
Gas Price = 2.00/gal	71.753	4.605	46.603
	(13.300)**	(11.011)	(30.426)
Gas Price = 3.00/gal	156.553	10.922	236.297
	(21.807)**	(15.438)	(55.240)**
Gas Price = 4.00/gal	241.353	17.238	425.992
	(33.967)**	(23.899)	(102.010)**

(d) Panel D: Bought Vehicle 4th GPM Quartile

	(1)	(2)	(3)
Gas Price = 2.00/gal	-58.920	-18.095	-39.611
	(9.487)**	(9.058)*	(20.474)
Gas Price = 3.00/gal	-124.111	-42.355	-170.980
	(15.808)**	(12.538)**	(37.907)**
Gas Price = 4.00/gal	-189.302	-66.614	-302.349
	(24.757)**	(19.569)**	(70.180)**

Marginal effect of the current gasoline price on the probability a household purchases a vehicle in the GPM quartile specified in table section header. Delta method standard errors robust to heteroskedasticity shown in parentheses. *,**,*** denote results significant at the 10%, 5%, and 1% levels, respectively.

Table 8: Net Effect of Kept Vehicle GPM Changes on Gasoline Consumption, New Vehicles

(a) 10% Reduction in Kept Vehicle GPM

	Observed Gasoline	Change in Gasoline Consumption (%Δ Kept vehicle gallons)		
Vehicle	Consumption (gal/yr)	Price Diff	Price DiD	Price Deviation DiD
Kept	537.64	-10.00	-10.00	-10.00
Bought	555.34	5.07	4.13	6.94
		(1.13)	(0.95)	(1.67)
Total	1,092.98	-4.93	-5.87	-3.06
		(1.13)	(0.95)	(1.67)

(b) 25% Reduction in Kept Vehicle GPM

	Observed Gasoline	Change in Gasoline Consumption (%Δ Kept vehicle gallons)		
Vehicle	Consumption (gal/yr)	Price Diff	Price DiD	Price Deviation DiD
Kept	537.64	-25.00	-25.00	-25.00
Bought	555.34	12.67	10.32	17.34
		(2.82)	(2.37)	(4.19)
Total	1,092.98	-12.33	-14.68	-7.66
		(2.82)	(2.37)	(4.19)

Predicted average change in fuel consumption resulting from an exogenous decrease in kept vehicle GPM of 10% (e.g., from 27.5 MPG to 30.6 MPG) or 25% (e.g., from 27.5 MPG to 36.7 MPG) for new vehicle purchases. Change in fuel economy expressed as percentage of annual gasoline consumption of the kept vehicle. Standard errors computed using the delta method shown in parentheses.

Table 9: Net Effect of Kept Vehicle GPM Changes on Gasoline Consumption, Used Vehicles

(a) 10% Reduction in Kept Vehicle GPM

Vehicle	Observed Gasoline Consumption (gal/yr)	Change in Gasoline Consumption (%Δ Kept vehicle gallons)		
		Price Diff	Price DiD	Price Deviation DiD
Kept	569.12	-10.00	-10.00	-10.00
Bought	537.29	6.18	6.15	6.66
		(1.42)	(2.12)	(2.28)
Total	1,106.41	-3.82	-3.85	-3.34
		(1.42)	(2.12)	(2.28)

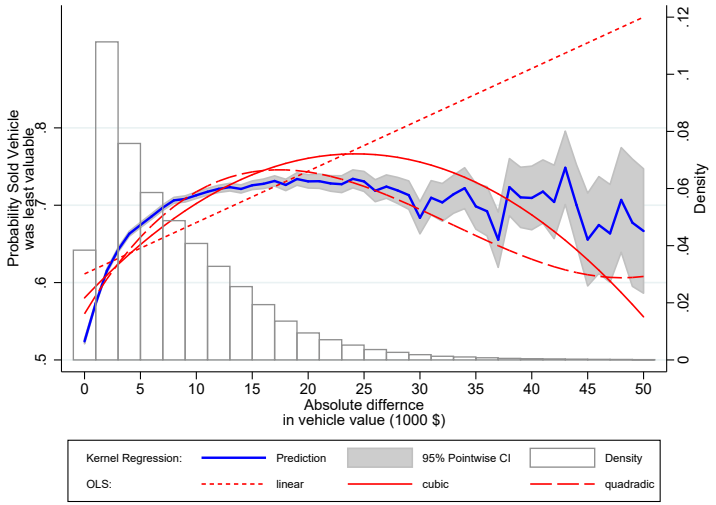
(b) 25% Reduction in Kept Vehicle GPM

Vehicle	Observed Gasoline Consumption (gal/yr)	Change in Gasoline Consumption (%Δ Kept vehicle gallons)		
		Price Diff	Price DiD	Price Deviation DiD
Kept	569.12	-25.00	-25.00	-25.00
Bought	537.29	15.44	15.38	16.64
		(3.55)	(5.29)	(5.69)
Total	1,106.41	-9.56	-9.62	-8.36
		(3.55)	(5.29)	(5.69)

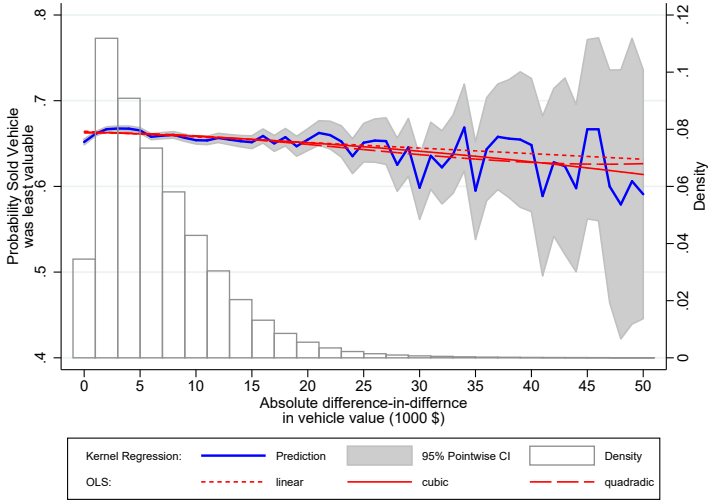
Predicted average change in fuel consumption resulting from an exogenous decrease in kept vehicle GPM of 10% (e.g., from 27.5 MPG to 30.6 MPG) or 25% (e.g., from 27.5 MPG to 36.7 MPG) for used vehicle purchases. Change in fuel economy expressed as percentage of annual gasoline consumption of the kept vehicle. Standard errors computed using the delta method shown in parentheses.

Figure 1: Instrumental Variables Reduced Form Relationships

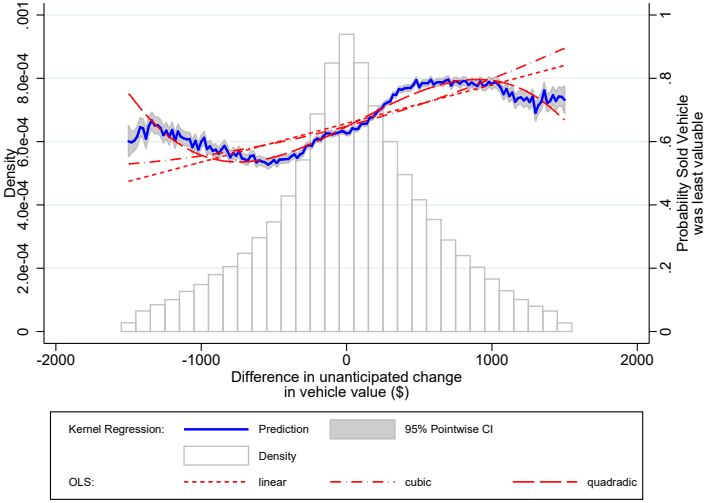
(a) Price Difference IV



(b) Price DiD IV



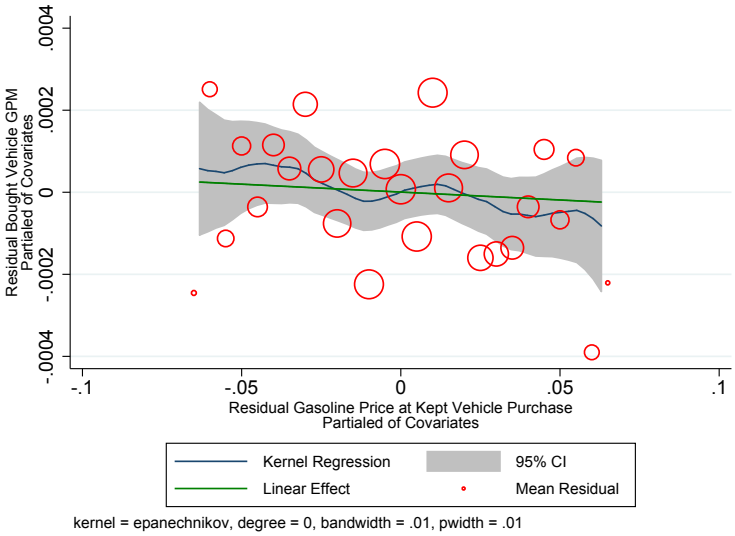
(c) Price Deviation DiD IV



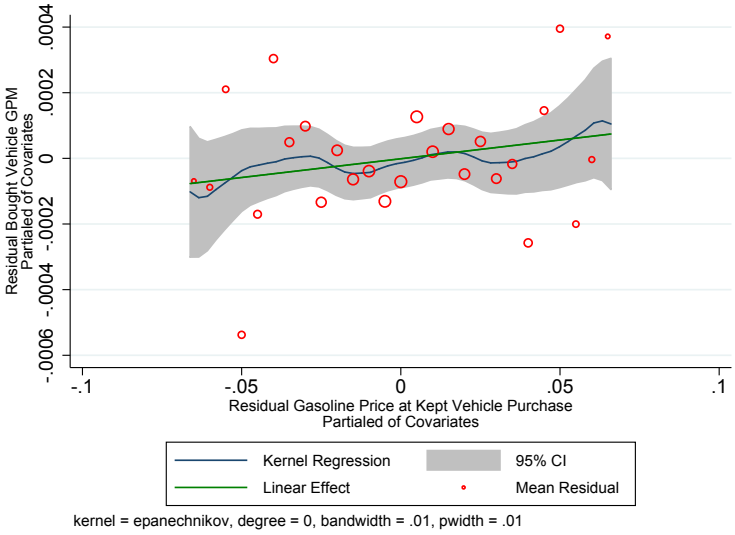
All 2x2 households. Probabilities conditional on a vehicle purchase (new or used) estimated within $1,000 bins. Binomial 95% confidence intervals shown in dashed lines. Values of the instruments in the Price Difference IV and Price DiD IV less than or greater than zero perfectly predict the least valuable vehicle in the portfolio and graphs are shown for the absolute value of these variables.

Figure 2: Reduced form relationship: Gas price at time of kept car purchase

(a) New Vehicle Purchases



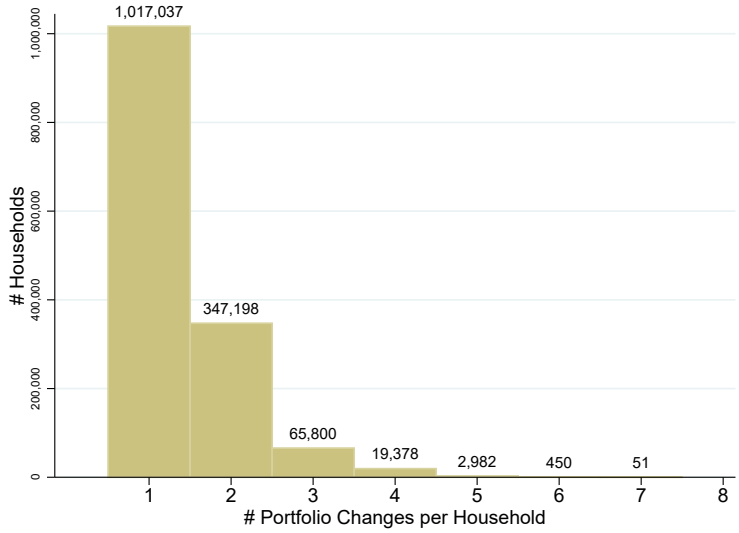
(b) Used Vehicle Purchases



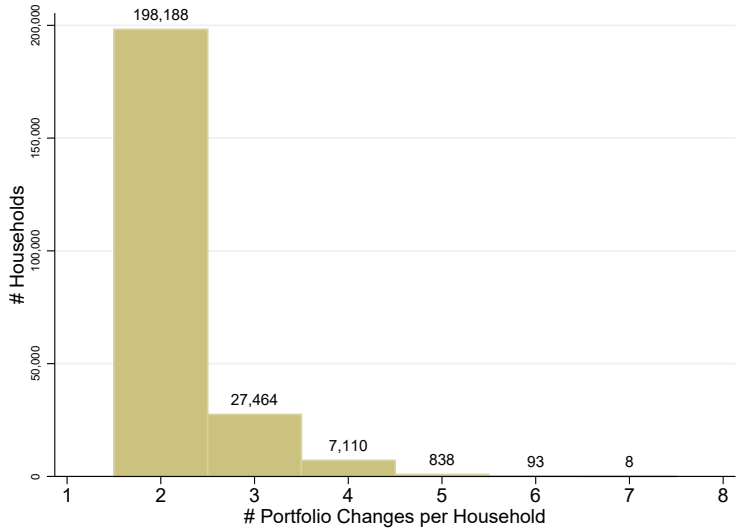
Plot of the reduced-form relationship between gasoline price at the time of kept vehicle purchase and the fuel economy (in GPM) of the purchased vehicle. Both variables are partialed of all other regression covariates. Graphs are limited to the 1st through 99th percentiles of residual kept vehicle gasoline price. Excludes observations where the household fixed effect perfectly predicts the outcome. Blue line is a kernel regression with Epanechnikov kernel and bandwidth 0.1. The gray band is the 95% confidence interval using the same kernel and bandwidth. The green line is the linear relationship estimated using OLS. Red circles are mean residuals for each 0.005 in kept vehicle GPM. The size of each circle is proportional to the number of observations used to compute the mean residual.

Figure 3: Number of Transactions per 2x2 Replacement Household

(a) Full Sample



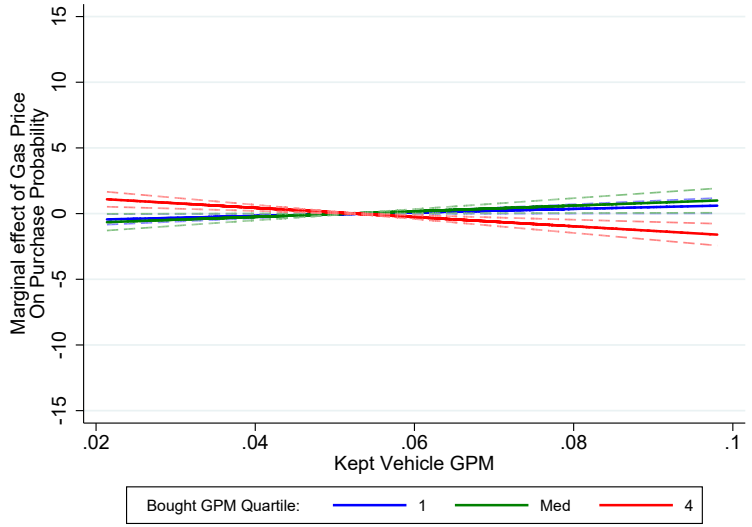
(b) IV Sample



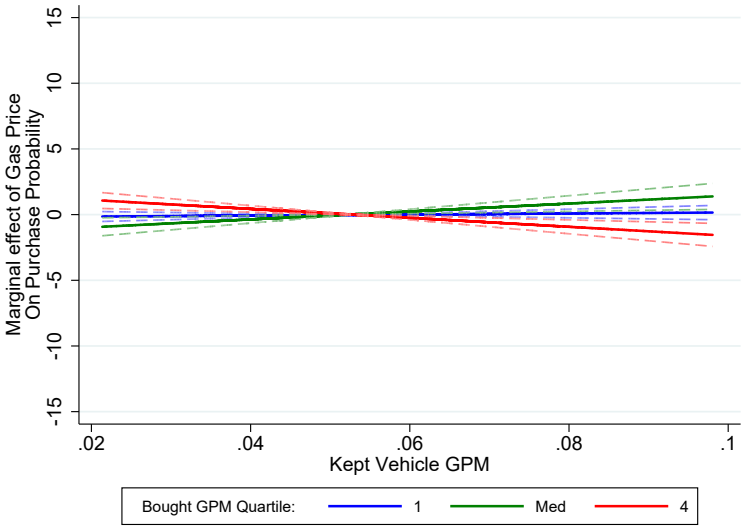
Distribution of the total number of observed vehicle transactions for each household from 2001 to 2007 for (a) the full sample of 2x2 replacement households and (b) households for which the data support deploying our IVs. In specifications including household fixed effects the fixed effect perfectly predicts the decision of a household if it only engages in one transaction. Other model parameters are identified by households engaging in multiple transactions from 2001 to 2007.

Figure 4: Marginal Effect of Gasoline Price, New Vehicles

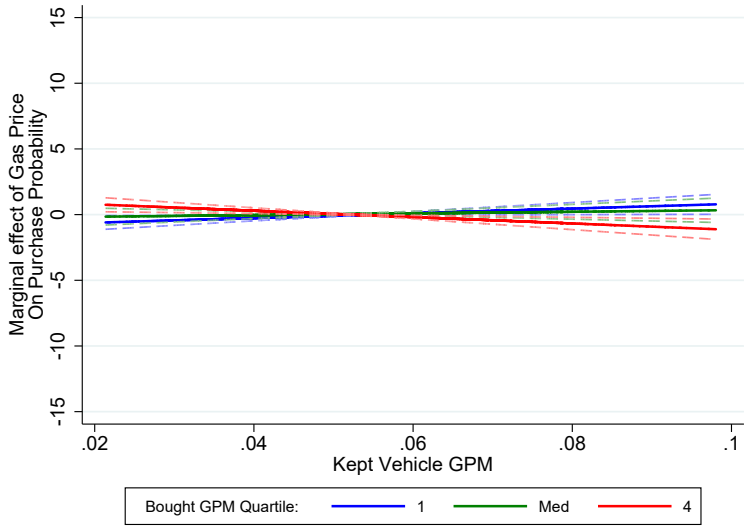
(a) Price Difference IV



(b) Price DiD IV



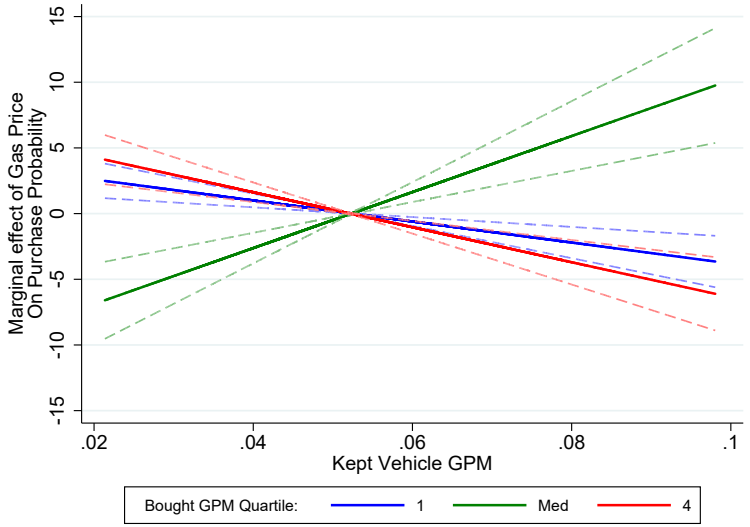
(c) Price Deviation DiD IV



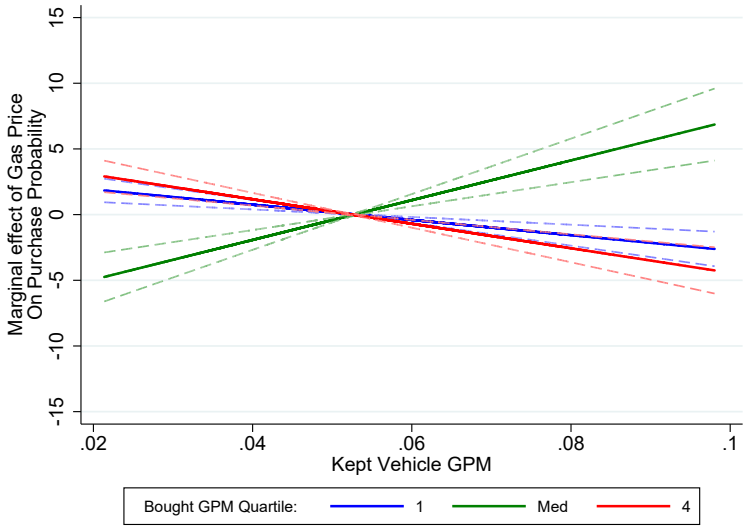
Pointwise 95% confidence intervals robust to heteroskedasticity shown in dashed lines.

Figure 5: Marginal Effect of Gasoline Price, Used Vehicles

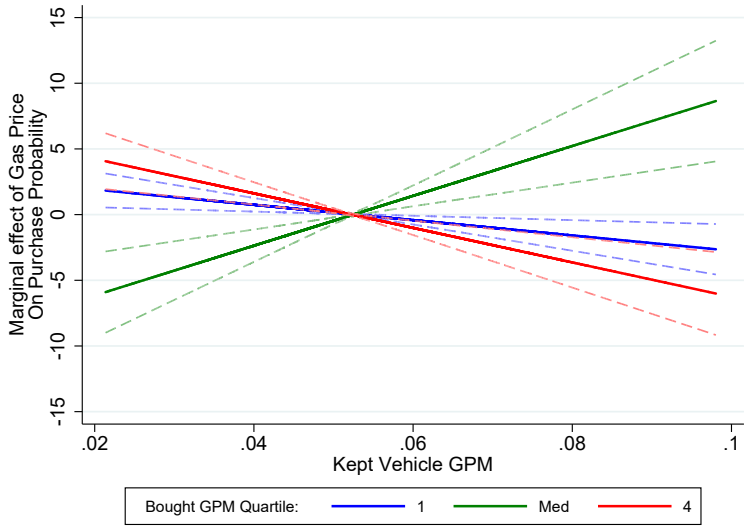
(a) Price Difference IV



(b) Price DiD IV



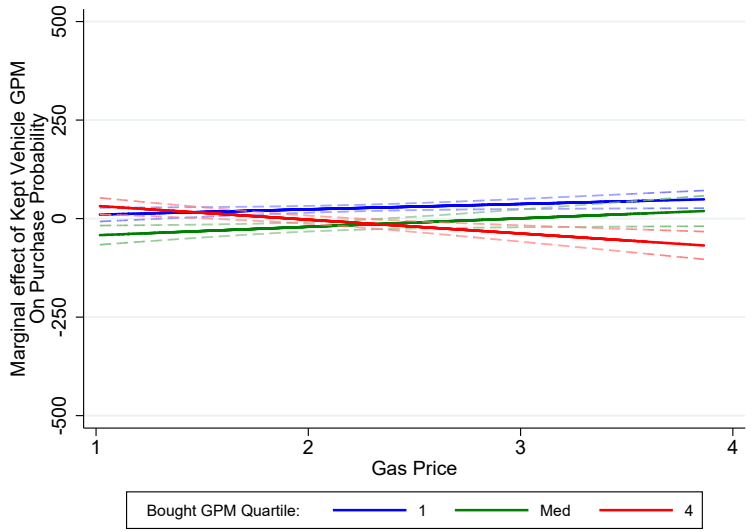
(c) Price Deviation DiD IV



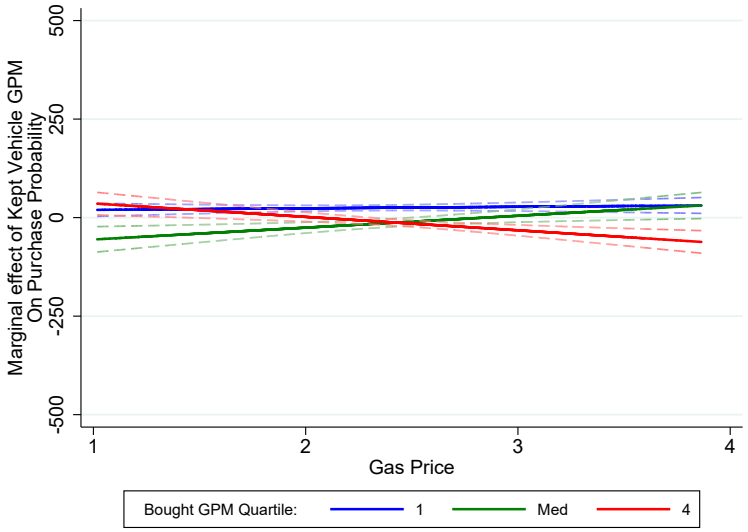
Pointwise 95% confidence intervals robust to heteroskedasticity shown in dashed lines.

Figure 6: Marginal Effect of Kept Vehicle GPM, New Vehicles

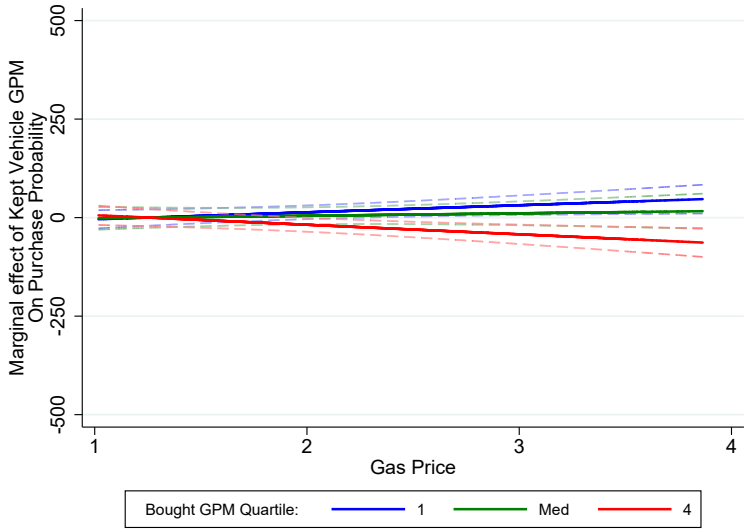
(a) Price Difference IV



(b) Price DiD IV



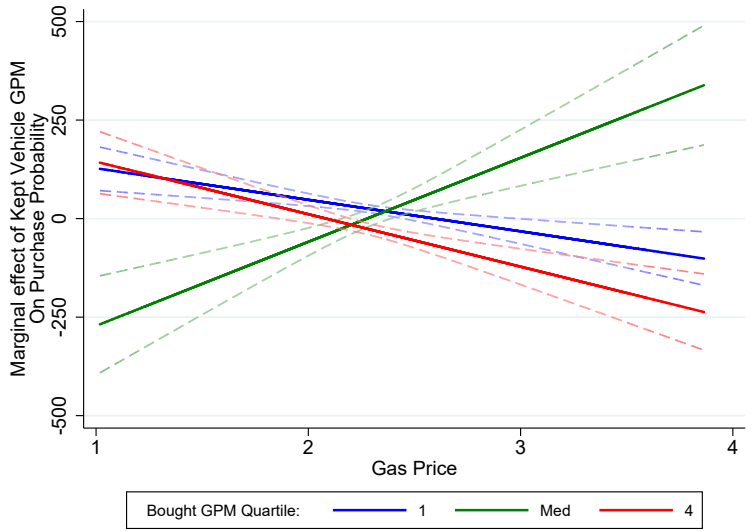
(c) Price Deviation DiD IV



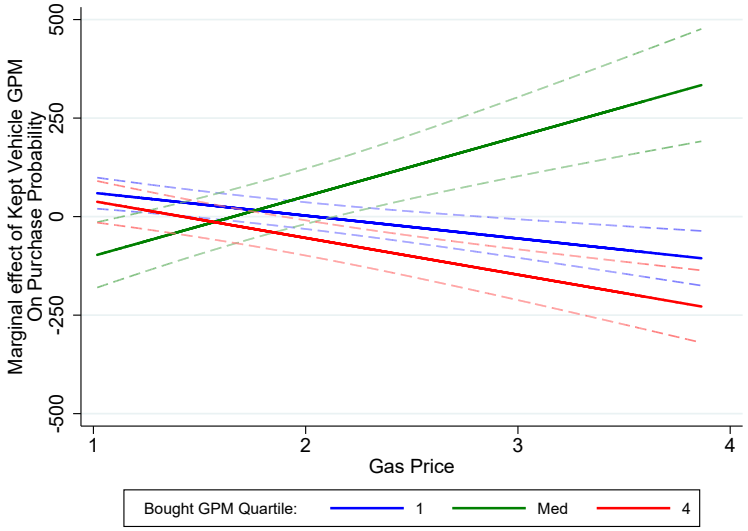
Pointwise 95% confidence intervals robust to heteroskedasticity shown in dashed lines.

Figure 7: Marginal Effect of Kept Vehicle GPM, Used Vehicles

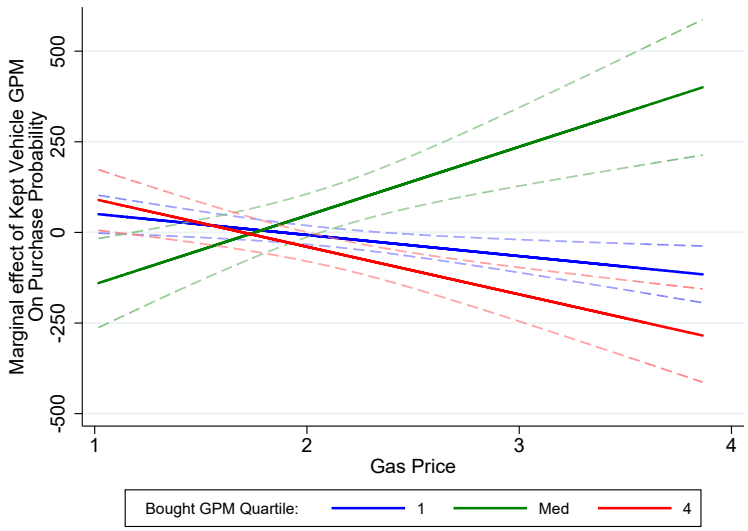
(a) Price Difference IV



(b) Price DiD IV



(c) Price Deviation DiD IV



Pointwise 95% confidence intervals robust to heteroskedasticity shown in dashed lines.