Heterogeneous (Mis-) Perceptions of Energy Costs: Implications for Measurement and Policy Design

Sébastien Houde,* Erica Myers†

October 25, 2018

Preliminary

Abstract

How consumers perceive different aspects of product cost, such as sales tax, shipping and handling charges, and energy operating expenses among others, have important welfare implications for policy. In this paper, we estimate heterogeneous perceptions of energy costs in the U.S. appliance market using a revealed preference approach. We recover a non-parametric distribution and show that while the largest share of consumers correctly perceives energy costs, a significant share undervalues them, and smaller shares either significantly overvalues or does not pay attention to them. These patterns are strikingly similar across income groups. We simulate the welfare effects of policies targeting externalities in the presence of heterogeneous misperceptions and show that standards largely outperform taxes. Standards' key advantage is that they reduce variance in energy operating, which ameliorates the distortionary effects from potential misperceptions.

JEL Codes: Q41, Q50, L15, D12, D83.

Key words: demand estimation, misperceptions, energy efficiency gap, behavioral welfare economics.

*ETH Zürich, shoude@ethz.ch.

†Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign, ecmyers@illinois.edu.

We would like to thank Meredith Fowlie, Christopher Knittle, Richard Newell, Ryan Kellogg, Frank Wolak, Thomas Colvert, Kochiro Ito, Anna Alberini, Akshaya Jha, Tomomichi Amano, Tatyana Deryugina, in addition of numerous seminar participants. The U.S. Department of Energy and the National Science Foundation provided funding for this project through the NBER Economics of Energy Market program. This research is also part of the activities of SCCER CREST, which is financially supported by the Swiss Commission for Technology and Innovation (CTI) / Innosuisse.

1. Introduction

Behavioral economists have long pointed out that limited cognition, biases, and heuristics have important implications for a wide range of economic decisions. However, it is only recently that empirical evidence of such behavior has been demonstrated in naturally-ocurring markets.[1] There is now a large and growing literature demonstrating that consumers are prone to making mistakes across various economically important markets, such as in choosing health care plans, mutual funds, mortgages, which foods to consume, and in accounting for sales taxes and various shrouded attributes like bank fees, mini bar fees, and shipping and handling expenses among others.[2] For policy design, it is crucial to understand what type of behavior can be attributable to mistakes and how to empirically detect and quantify them. For example, taxation, one of the most commonly used policy levers, has dramatically different effects on behavior if consumers misperceive aspects of product costs targeted by taxes than if they do not.[3] Recent work in behavioral public finance has shown that the presence of misperceptions, in particular heterogeneous misperceptions, is a critical input for optimal policy design. It influences the optimal level of a particular policy instrument, but also can determine the type and/or combination of instruments that should be used (Farhi and Gabaix 2018; Allcott et al. 2014).

While characterizing misperceptions can be crucial for policy design, it is challenging in practice. The researcher needs: 1) a rich description of consumer behavior in the "naturally occurring" environment, where they make mistakes, 2) a rich description of consumer behavior in the "welfare-relevant" environment, where they do not make mistakes, and 3) the mapping between them (Bernheim and Taubinsky 2018). Previous work has relied on two approaches to measure misperceptions. The first approach is to use "artefactual field experiments" (Harrison and List

[1] The series "Anomalies" by Richard Thaler and his co-authors published in the Journal of Economic Perspectives (1987-2006) is probably the first effort to report real-world manifestations of biases and heuristics that behavioral economists previously documented in laboratory settings.

[2] See for example mistakes in healthcare plan choice (Abaluck and Gruber 2011; Kling et al. 2012; Handel and Kolstad 2015; Heiss et al. 2016; Ketcham et al. 2016a), mutual funds (Barber et al. 2005), mortgages (Allen et al. 2014; Guiso et al. 2018), schools (Jensen 2010), and nutrition (Bollinger et al. 2011). In addition see the following for inattention to sales taxes (Chetty et al. 2009), Taubinsky and Rees-Jones (2018) and other shrouded attributes (Ellison 2005; Gabaix and Laibson 2006; Ellison and Ellison 2009; Hossain and Morgan 2006).

[3] In this paper, we will use the term "misperception" to refer to any biases, heuristics, or biased beliefs that induce consumers to undervalue/overvalue an aspect of product cost.

2004) to create linkages between the naturally occurring and welfare-relevant environment with experimental choice environments (Allcott and Taubinsky 2015; Taubinsky and Rees-Jones 2018). The second approach is to estimate a "behavioral econometric model," (DellaVigna 2018) which explicitly characterizes a particular departure from the neo-classical rational model, then simulate choices where the behavioral departure is turned off to create a mapping between decision utility and the welfare-relevant experienced utility.

In this paper, we develop a revealed preference estimator to recover heterogeneity in perceptions for one dimension of product costs: energy operating costs. This is the first attempt that we are aware of—not only in the energy context, but more broadly—to use a revealed preference approach that explicitly focuses on recovering heterogeneous (mis-) perceptions. Our semi-parametric estimator recovers a non-parametric distribution of perceptions of energy costs. The advantage of our approach is that it allows us to depict heterogeneity in preferences, without requiring us to take a stand on rational behavior. Post-estimation, different analysts can then use different assumptions to judge if the distribution of preferences is consistent with a fully rational model or not.

We apply our estimator to the U.S. appliance market, in particular to refrigerator purchase decisions. Characterizing consumer perception of energy costs is crucially important for designing successful policy instruments for reducing greenhouse gas (GHG) emissions. Whether or not a particular level of an energy tax or appliance standard will be cost-effective depends on how consumers respond to the policy. Recent work has found modest or no undervaluation of energy costs on average in the context of cars, housing, and appliances (Busse et al. 2013; Allcott and Wozny 2014; Grigolon et al. 2018; Sallee et al. 2016; Myers 2018; Houde and Myers 2018). In this paper, we show that the average masks substantial heterogeneity and can be a misleading statistic for policy design. We find that while the largest share of consumers appears to correctly perceive energy costs, a significant share undervalues them, and smaller shares either significantly overvalues or does not pay attention to them. These patterns are strikingly similar across income groups and robust to various normative assumptions used to determine the true experienced utility of consumers.

Underlying the modest level of misperception that we found on average are heterogeneity patterns that suggest that a proportion of consumers are prone to making mistakes, either severely over or undervaluing energy costs. We support this conclusion by providing additional empirical evidence on two particular mechanisms that can explain these types of mistakes: inattention and biased beliefs. We examine the role of inattention with a fully non-parametric estimator that exploits the existence of pairs of identical refrigerator models that differ only with respect to energy efficiency and price. While the more efficient model is almost always more expensive, there are sales events when it is less expensive, dominant option. We find that 1/3 to 1/2 of consumers still buy the dominated option, a level of inattention that could rationalize high levels of undervaluation of energy costs. Second we assess the role of biased beliefs using a survey designed to assess energy literacy, we also find that a significant share of consumers have biased beliefs, which can rationalize the severe misperception we observe particularly in overvaluing energy costs.

We show how heterogeneous misperceptions can impact the design of Pigouvian policies and explore the tradeoffs of taxes versus efficiency standards in the presence of potential misperceptions. Our framework builds on the work of Farhi and Gabaix (2018) and Houde and Aldy (2017a), deriving expressions for the optimal *behavioral* policies that aim to address negative externalities while accounting for misperceptions. We show that while heterogeneity in perception can significantly impact the level of optimal price instruments, it has small and possibly no effect on the design of quantity instruments. When we use our estimated demand model to compare different types of standards and Pigouvian taxes, we find that standards largely outperform taxes. We show that standards, unlike taxes, can internalize externalities and misperceptions at once. The key advantage of standards is that they may reduce variance in energy operating costs in the choice set, which reduces the distortionary effects of misperception in trading off energy costs with other attributes. With low variance in operating costs in the choice set, misperceptions of energy costs have little, and in some special cases no, effects on substitution patterns and ultimately welfare.

The empirical strategy we develop to recover unobserved heterogeneity non-parametrically is well-suited for Big Data demand analysis in other contexts like ours, where the number of observations and products are both high dimensional. We first show that focusing on estimating non-parametric distribution is crucial in our setting. Using Monte Carlo evidence, we demonstrate that relying on parametric assumptions to specify the distribution of heterogeneity is likely to yield large biases about the true nature of the heterogeneity presents in a market. This finding is particularly important given that characterizing heterogeneity is the main objective of this paper.

We then propose a two-step estimator, which we adapt from Fox, Kim, Ryan, and Bajari (2011) (FKRB, thereafter) to recover a multi-dimensional non-parametric distribution of preferences while addressing the curse of dimensionality. The FKRB estimator has been used in few applications (e.g., Nevo et al. 2016; Blundell et al. 2018). To our knowledge, we are the first to use it for a demand analysis in a context where there are several hundred of options in the choice set and there is a need to account for a large number of product fixed effects. We also provide Monte Carlo evidence that our estimator can achieve consistency in datasets typically used by applied researchers. For instance, we use administrative data and perform the estimation with a sample of close to 200,000 consumers and a choice set that spans more than 400 options.

The remainder of the paper is organized as follows. In the next section, we discuss the data and choice environment for our empirical investigation. In Section 3, we present our empirical framework for recovering heterogenous perceptions. In Section 4, we present the results of our estimation and in Section 5 we examine the role of inattention and beliefs in driving the heterogeneity patterns that we observe. Section 6 follows and investigates the design and evaluation of behavioral environmental policies.

2. **Data and Environment**

Our empirical investigation focuses on the U.S. refrigerator market, which offers several advantages. First and foremost, refrigerators are one of the few appliance categories that consume a large amount of energy and have little variation in utilization across consumers. Although refrigerator energy costs could be subject to idiosyncratic variation across households, the characteristics of a refrigerator such as its size, door design, and presence of ice maker are the main determinants of its energy costs. Therefore, it is not necessary to explicitly model the endogeneity of the utilization and purchase decisions, which simplifies the estimation. Second, the U.S. refrigerator market is subject to rich variation in refrigerator prices, energy costs, rebates for energy efficient appliances, and choice sets that allow us to identify the preference parameters and infer the degree of misperception to energy costs. Third, refrigerators is an important market in the U.S. and elsewhere, which is expected to grow particularly fast in developing countries in the upcoming decades (Gertler et al.

2016). Contributing to the design of policies that improve the energy efficiency of refrigerators is thus important to reduce the negative externalities associated with household energy use.

The main data source used for the estimation consists of transaction level data from a large U.S. appliance retailer. The sample includes all transactions where a refrigerator was purchased during the period 2007-2012. We observe each transaction, which contains information about the price paid by the consumer, the zip code of the store where the purchase was made, the manufacturer model number of the model purchased, and a transaction identifier that tracks consumers making multiple purchases. For a large subset of transactions, the identifier is matched with household demographics collected by a data aggregator (Table 1). Detailed attribute information for each manufacturer model number is also available and includes: manufacturers' reported energy use, dimensions (width, height, depth), whether a product is certified Energy Star, the presence of ice maker, color, brand, door design, and several other features pertaining to design and technology options.

One particular feature of the U.S. appliance market is that appliance retailers, such as ours, have a national price policy and retail prices are subject to large and frequent variations. The price of each refrigerator model at this same retailer is subject to weekly variation that can exceed 20% and that variation is model-specific and not perfectly correlated across brands. Figure 1 displays an example of this variation in weekly price over the study period for the top 9 sales ranked models from a particular refrigerator brand. The red line shows the median weekly price change across all zip codes with the 25th and 75th percentile depicted with gray bands. The blue line shows the remaining variation in weekly price change after controlling for week-of-sample fixed effects interacted with brand dummies. Even with these controls, large variation in in price remains suggesting that the model-specific variation we observe is highly idiosyncratic and is generated by the retailer's dynamic pricing algorithm. We exploit this variation to identify the coefficient on price.

We match the transaction data with local energy prices and rebate information. Energy prices are constructed from the form 861 of the Energy Information Administration (EIA), which contains revenue and quantity of kWh consumed by residential consumers. Together, these variables provide a measure of average electricity price for each electric utility operating in the U.S. The EIA also provides information about which utility is operating in each county, which allows us to compute average electricity prices at the county level. If more than one utility serves a county, we take the average of those utilities' prices. Figures 2 gives a sense of the variation in electricity price across regions and time. It depicts mean sales-weighted annual electricity price for each state in a U.S. census division. Each line represents the sales weighted average price for a state plotted for each of the 9 U.S. census divisions. Prices vary quite a bit regionally, with the highest prices in New England and the lowest prices in the Midwest and South. There is also variation over time in price with some states experiencing price increases and others price decreases over the study period.

We estimate the annual energy costs for each model-year-store location by multiplying the annual kwh consumption reported by the manufacturer by the energy price in the store's county. Figure 3 displays density plots of model's prices and lifetime energy costs. The first panel plots the distribution of prices paid. Almost all models sold in our sample are less than $2000, though there are some much higher priced models available. The second panel shows how the the distribution of life time energy cost of the models sold varies for the 10th, 50th, and 90th percentile of electricity price. Using an expected lifetime of 18 years and a 5% discount rate the mean lifetime costs range from $555 for the 10th percentile of energy price to $1000 for the 90th percentile. The third panel shows the distribution of the ratio of lifetime cost to purchase price, with means ranging from .44 at the 10th percentile of energy price to .79 at the 90th percentile. This variation in energy costs for particular models across time and space identify the coefficient on energy costs.

Rebates for energy efficient appliances were offered during the sample period by both state governments and electric utilities. The State Energy Efficiency Appliance Rebate Program (SEEARP) was funded as part of the stimulus package of the American Recovery Act. This program led to generous rebates for Energy Star certified products during the year 2010 and 2011. Several electric utilities also offer rebates for Energy Star certified-refrigerators. Both of the rebate programs vary across time and regions. Houde and Aldy (2017b) find that consumers responded to SEEARP, but a large fraction of consumers were inframarginal to the program. Houde and Aldy (2017) and Datta and Filippini (2016) also find that rebates offered by electric utilities have a modest impact on demand due to a low take-up rate.

We carry the estimation using a large subsample of transactions. Each model that we estimate uses a random subsample of approximately 32,000 transactions. We restrict the sample to transactions made by households owning their housing unit[4] with the goal of focusing on transactions made by consumers who are likely to pay the energy operating costs of their appliances.

3. Recovering Heterogeneous Misperceptions

3.1. Framework

In this section, we present the standard framework that has been used to quantify the average degree of misperception of energy operating costs. We then discuss how to apply this framework to recover heterogeneous misperceptions and the challenges in doing so.

Setup. Much of the literature estimating demand for energy using durable uses a discrete choice framework where consumer i chooses among J different alternatives. The two main variables to consider in the purchasing decision is the price (capital cost) of product j, denoted P_j, and the future energy operating costs over the entire expected lifetime of the product, denoted E_j. For now, we abstract away from uncertainty and consumer-specific heterogeneity in the product lifetime, future energy prices, utilization, depreciation, and discount factor for ease of exposition.[5] We simply assume that E_j is the exact measure of expected energy operating costs discounted with a normal rate of return. Consumer i values product j as follows:

$$U_{ij} = \gamma_j - \eta_i P_j - \theta_i E_j + \epsilon_{ij} \tag{1}$$

where γ_j is the vertical quality of product j and captures all the attributes of the products and ϵ_{ij} captures idiosyncratic preferences. Numerous studies have estimated variants of this model with the goal of identifying the preference parameters for price and energy cost: η and θ, respectively. The

[4]The data do not explicitly identify transactions that are made by households. We infer this information using a transaction identifier that tracks multiple purchases of customers. We classify customers that purchase more than two refrigerators during the period 2008-2012 as non-households. This criterion is a conservative way to rule out contractors and other entities that buy a large number of appliances in bulk.

[5]We explain the assumptions we make and the heterogeneity in discount factor that we explore in what follows.

common approach used to quantify the degree of misperception is to test whether consumers trade off one dollar of energy costs the same way they trade off one dollar of purchase price. Formally, the test of misperception is whether $\theta/\eta \neq 1$. An alternative, but equivalent test that has been widely used in the literature solves for the implicit discount rate such that the ratio $\theta/\eta = 1$. Using this latter approach, an implicit discount rate that is markedly above the normal rate of return is seen as a sign that consumers undervalue energy costs.

Average Misperception. If the preference parameters for prices and energy costs η and θ vary in the population, the statistic used to measure the average degree of misperception: $m = \theta/\eta$, is the ratio of two random variables. A closed form solution for the distribution of m exists only for a few specific distributions (e.g., two lognormals), but the first moment of that distribution can be easily approximated given any distributions of η and θ, as we show below.

Proposition 1. *The first-order Taylor approximation of $E[m] = E[\theta/\eta] \approx E[\theta]/E[\eta]$. The second-order Taylor approximation of $E[m]$ is:*

$$\frac{E[\theta]}{E[\eta]} - \frac{cov(\eta, \theta)}{E[\eta]^2} + \frac{Var(\eta)E[\theta]}{E[\eta]^3}$$

Proposition 1 shows that, even if one is only interested in the average degree of misperception, heterogeneous responses to both prices and energy costs and how these responses are correlated must be considered.[6]

Heterogeneous Misperceptions. To recover heterogeneity in misperceptions, Proposition 1 can be extended to approximate the higher moments of the distributions. To circumvent the bias induces by the approximation, a simple solution is to estimate a discrete choice model in the willingness pay dimension (Revelt and Train 1998). This estimation procedure requires a simple transformation of the estimating equation and no additional computations relative to the standard approach used in the discrete choice literature. To implement the estimation in the willingness pay dimension, it suffices to transform Equation 1 as follows:

[6]Note that several studies that have quantified misperceptions of energy costs only report the first-order Taylor approximation shown in Proposition 1. This approximation induces a bias of the order of:

$$-\frac{Cov[\eta, \theta]}{E[\eta]^2} + \frac{Var[\eta] \cdot E[\theta]}{E[\eta]^3} + \mathcal{O}(n^{-1}) \tag{2}$$

$$U_{ij} = \eta_i(P_j + m_i E_j) + \gamma_j + \epsilon_{ij} \tag{3}$$

and estimate η_i and m_i. The advantage of this approach is that the estimation directly yields an estimate of $E[m]$, or a distribution of m if it is specified as a random parameter. No further post-estimation calculations and approximations are thus required. For an estimation standpoint, recovering the exact average degree of misperception, $E[m]$, can thus be straightforward. There are, however, challenges in estimating a distribution of misperceptions, especially in the context of energy-using durables.

The main challenge arises when misperceptions are treated as unobserved heterogeneity. Given that households purchase energy-using durables infrequently, this means that unobserved heterogeneity must be recovered using only one choice situation per household. In this context, estimating the true distribution of heterogeneity is particularly challenging. For instance, in a Monte Carlo experiment, Train (2009) shows that with one choice situation, the mixed logit model can recover only approximately 40% of the variation in population preferences, but with 10 choice situations, more than 80% of the variation is captured.

In our own Monte Carlo experiment (Appendix A2), we reach a similar conclusion regarding the mixed logit and its ability to recover higher moments of the true distribution of preferences when the data have only one choice situation per household. We investigate the role of the number of observations (i.e., decision-makers), the number of alternatives, and mis-specification in the parametric distribution of preferences. Across the different scenarios, we find that the mixed logit performs well at recovering the mean values of the true population parameters. However, the estimate of the covariance matrix of the distribution of preferences are imprecise and subject to large biases. Scenarios where the parametric distribution is mis-specified are particularly discouraging. For instance, if the true distribution is a mixture of normals, but we specify that preferences follow a unimodal normal, the estimated covariance matrix would suggest a much wider distribution of preferences compared to the true data generating process.

Given that we are interested in the higher moments of the distributions of misperceptions, flexibly recovering unobserved heterogeneity is crucial. In our empirical application, we adapt the semi-parametric estimator of FKRB, which allows us to estimate a non-parametric distribution of m. We then extend our Monte Carlo experiment and show that it performs well in recovering complex distributions of unobserved heterogeneity.

3.2. **Estimation**

The intuition of the FKRB estimator is that a continuous distribution of random parameters can be approximated by a discrete distribution defined over the discretization of the support of the continuous distribution. This approximation yields a simple estimator that can be implemented as follows. First, discretize the support of the random parameters into a large number of grid points (say K). Second, evaluate a (parametric) choice model at each grid point $k \in K$. Third, integrate over the discrete distribution by simply weighting and summing the choice model at each grid point. The estimator returns the weights, which correspond to the discrete probability density function (pdf) of the random parameters. This estimated discrete pdf is a fully non-parametric distribution.

To illustrate, consider our setting where we are interested in evaluating the joint distribution of $F(\eta, m)$. We first discretize the support of η and m into K grid points: $\beta^k = \{\eta^k, m^k\}$, $k \in K$. We then compute the choice model for each β^k using a parametric model where the probability of choosing product j given β^k is noted $P_j(\beta^k) = P_j^k$. For the parametric model, we are using the conditional logit. The choice probability, P_j, is then a mixture of K conditional logit models:

$$P_j = \int P_j(\eta, m) dF(\eta, m) \approx \sum_k^K \alpha^k P_j^k \tag{4}$$

where $\sum_k^K \alpha^k = 1$ because the weights are a discrete probability density function that approximate the true underlying continuous function. By choosing a parametric form for the choice model, each P_j^k can be first be computed for each grid point and then treated as data in the estimation. The estimator is thus semi-parametric and the estimation can simply proceed by running a linear regression with P_j as the dependent variable, P_j^k as regressors, and α^k, $\forall k \in K$ as coefficients to be estimated. To ensure that the weights α^k sum to one, constrained linear least squares must be used with the constraint: $\sum_k^K \alpha^k = 1$.

The FKRB estimator is appealing for a number of reasons. First, it is a computationally simple way to recover a fully non-parametric distribution of heterogeneity. Second, constrained linear least squares is guaranteed to provide a global optimum if a solution exists, a non-trivial advantage over other estimators of random coefficients models, such as maximum likelihood, the EM algorithm, GMM, or bayesian methods, which are all prone to local optima and convergence issues.

The main weakness of the FKRB estimator is that it suffers from the curse of dimensionality. The number of grid points increases exponentially with the number of random coefficients. For instance, a model with three random coefficients discretized with 100 grid points in each dimension must be evaluated for a combination of 100^3 grid points. If the number of preference parameters is large, it becomes rapidly intractable to model each parameter with a random coefficient. FKRB propose two solutions. First, some parameters can be estimated in a first stage and be treated as data when estimating the weights α_k in a second stage. For instance, a simple conditional logit could be estimated for the first stage, and then the FKRB can be implemented where only a subset of the parameters are treated as heterogeneous. This approach is particularly appealing with models with a large number of fixed effects. It should, however, be noted that this two-step estimation may not produce consistent estimates and standard errors need to be adjusted in the second step. The second approach proposed by FKRB is to estimate the model in one step with non-linear constrained least-squares. The feasibility of this estimator is context specific. The non-linear optimization is also not guaranteed to converge to a global optimum.

In our application, the curse of dimensionality is an important issue given that we have a high dimensional parameter space due to the large number of fixed effects. Moreover, the choice probabilities take a relatively long time to compute given the large number of observations used for the estimation. We thus favor the two-step approach. To investigate the consistency of this approach, we extended our Monte Carlo experiment. The main question in implementing the two-step approach is which model to use in the first-stage, and in particular if a mis-specified first-stage as implications for the estimation of the non-parametric distribution in the second-stage.

We compare two mis-specified first-stage approaches: (1) a conditional logit model and (2) a parametric mixed logit model with a normal distribution. The conditional logit, a homogenous estimator, will be mis-specified given that the true data generating process is subject to preference heterogeneity. The mixed logit model is also possibly mis-specified if the parametric distribution does not correspond to the true distribution of preferences. However, relative to the conditional logit, the mis-specification is of different nature, concerning only the shape of the distribution.

The advantage of using an homogeneous model in the first-stage is that it is computationally easy and fast to implement. The estimates of the fixed effects, however, are likely to be subject to biases if some of the true population parameters, e.g., η and m, are random. The parametric mixed logit, although, more computationally demanding would help correcting for this bias—as we showed earlier, a mixed logit, even if mis-specified, still performs well in recovering the means of the true population parameter estimates. To obtain consistency with the FKRB estimator, we should strive to evaluate the model at the true fixed effects values, which a mis-specified mixed logit should be able to deliver. We confirm this intuition in our Monte Carlo experiment (Appendix A2). We show that the FKRB estimator implemented with a two-step approach performs well when the fixed effects are first estimated mixed logit. This is true even when the distribution of the mixed logit is mis-specified. On the other hand, simply using the fixed effects estimates from the conditional logit tends to lead to a downward bias in the estimated non-parametric distribution, especially if the distribution of preferences follows a complex distribution.

Based on these insights, we implement the FKRB estimator as follows. The parametric choice model for each grid point $k \in K$ is the conditional logit defined in the willingness to pay dimension with alternative-specific utility given by

$$U_{ijrt}^{k} = \eta^{k}(P_{jrt} + m^{k}E_{jrt}) + \tau ES_{jt} + \phi Rebate_{jrt} + \gamma_j + Demo_i \times Att_t + \epsilon_{ijrt} \tag{5}$$

The model has a parsimonious set of controls: the Energy Star certification (ES_{jt}), Energy Star rebates ($Rebate_{jrt}$), product fixed effects (γ_j), and demographic information interacted with a subset of attributes ($Demo_i \times Att_t$). As we show below, the estimates from conditional logit are robust to various set of controls. This parsimonious specification is thus sufficient to not induce a large bias, while being computationally tractable.

Only the parameters η and m are random coefficients. In the first step, we estimate all coefficients with a mixed logit. In the second step, we fix all the regressors, except η and m. The grid points for η and m are defining by a large grid around the estimates of the mixed logit. For one robustness test, we also treat τ as being heterogeneous.

For each transaction, we infer a zip code-trimester-specific choice set, i.e., all models offered in a given zip code and during a trimester are considered to be in the consideration set of a consumer.[7] We focus on modelling the purchase decision conditional on the fact that a consumer has decided to buy a refrigerator at time t and in location r. The timing decision and the choice of the retail chain store are thus not explicitly modelled. The marginal effects of the coefficients on price and energy costs thus capture the substitution across different models offered.

To recover the standard error in the second-stage and to account for uncertainty due to the two-step estimation, we implement the subsampling bootstrap (). This consists of slicing the dataset in S subsamples, without replacement, and performing the estimation for each subsample. The estimated joint pdfs are averaged across all subsamples. The standard errors of each discrete weight of the pdf capture the variation across subsamples. When applying the subsampling bootstrap, we should be careful in making inference for parameters situated at the edge of the support of the distribution, which often arise with the FKRB' estimator. The choice of the grid points, especially the end of the support should be done with caution. Therefore, we report robustness tests investigating a denser grid and a smaller grid. We provide additional details about the estimation procedure in Appendix A3.

In addition, to investigate heterogeneity in the distribution of misperceptions across income, we estimate the joint distribution of η and m separately for six different income groups: <$30k, $30k-$50k, $50k-$75k, $75k-$100k, $100k-$150k, and >$150k. We implement the estimator on six different samples using the same approach described above. Each sample is drawn from the universe of transactions made by households in a particular income group.

[7] ##% of zip codes have only one store. Our choice set are thus mostly store specific. We do not observe floor inventory. Therefore, a model is deemed to be offered if we observe at least one sale of that model at a given location and time period.

3.3. **Lifetime Energy Costs**

Quantifying misperceptions requires making assumptions about consumers' beliefs about future fuel prices, the lifespan of the refrigerator, and the discount rate. In this section, we describe the assumptions made in our analysis. Though given our estimation, analysts could apply different assumptions ex post and evaluate their effects on the distribution of misperceptions. For our analysis, we assume that consumers believe that annual electricity prices follow a no-change forecast, so that contemporaneous prices are the best predictor of future prices. In addition, we assume the average life expectancy for refrigerators of 18 years for all models and the average discount rate used by the DOE for appliance standards of 5% (the market rate of return) for all consumers.

As a robustness test, we estimate the sensitivity of the welfare estimates from our policy scenarios to our homogenous discount rate assumption. We explore a second scenario where we allow heterogenous discount rates to explain a wide range of variation in energy cost perceptions. We assume that any implied discount rate between 2% and 12% to be "rational," that is, we set $m = 1$. The return on 3 year U.S. Treasury bonds is close to 2% during our sample period, representing market returns from a risk averse investment strategy. On the other extreme, the average APR rate for credit cards during our period is around 12%, representing a cost of funds for consumers carrying credit card debt. For implied discount rates lower than 2%, we use a 2% discount rate to calculate misperception and for implied discount rates higher than 12%, we use a 12% discount rate to calculate misperception. We describe our results under these two scenarios in the next section.

4. **Results**

4.1. **Conditional Logit**

Table 2 presents the results for a simple conditional logit, which we use as a benchmark to compare to the model with heterogeneity. We first present a very parsimonious model where we only control for product fixed effects, the Energy Star certification status, and Energy Star rebates (Specification I). We then include interaction terms between demographic and product attributes (Specification II). In Specification III, we add county-fixed effects interacted with the ES status. Finally, we add brand-week fixed effects (Specification IV). Across these various specifications, the conditional logit is very robust: the average degree of misperception ranges from $m = 0.71$ to $m = 0.76$ if we consider that the correct rate of return is 5%. These estimates suggest a moderate undervaluation of energy costs, on average. In Houde and Myers (2018), we present alternative estimates based on various reduced-form estimators. Compared to these estimates, the conditional logit represents a lower bound on the estimate of $E[m]$.

4.2. **Non-Parametric Distribution of Misperceptions**

We first present the results of the FKRB estimator graphically. Figure 5 shows the joint pdf of η and m (Panel a), the marginal pdf of η (Panel b), and the marginal pdf of m (Panel c). In each panel, each pdf weight corresponds to an average of the weights estimated across the bootstrap iterations.[8] The uncertainty of each estimated weight is depicted by a circle or a star, where a star corresponds to an estimate that has a t-statistic of 1.1 or larger.[9] Given that we use sixteen slices for the subsampling bootstrap, the value 1.1 corresponds to the critical value of a one-tail t-test with ten degrees of freedom and a significance level of approximately 15%.

Focusing on the joint pdf (panel a), the size and the color of the markers represent the value of the pdf. We observe several estimates that have a pdf weight of 15% or above (the larger red markers), and several of these estimates tend to be statistically significant (i.e., they have a t-statistic greater than 1.1). Several estimates in the 0 to 15% range (blue and green markers) are also statistically significant. We also observe that the joint distribution is centered around the mean of η and m (depicted by the two black lines), but there are also several smaller modes located far from the mean values.

The distribution of misperceptions, m, is especially dispersed. The marginal pdf of m (panel b) has several modes and statistically significant weights at grid points located between zero and one. However, there are also statistically significant weights above one and below zero. A value of m

[8] The pdf displayed in each panel is effectively a point-by-point average of sixteen different pdfs. Therefore, the sum of the weights on Figure 5 may not sum to 1.

[9] We favor this approach to show uncertainty in the estimates over confidence intervals because the bootstrap procedure does not allow to recover confidence intervals for all estimated weights. This is due to the fact that a positive pdf weight for a particular grid point is sometimes found for only one of the sixteen bootstrap iterations. When this is case, we have only one degree of freedom and we cannot compute the standard error. In the figure, these estimates are coded as having a t-statistic below 1.1, although the statistic is effectively undefined.

greater than one suggests that there is share of consumers that place a very high weight on energy operating costs in their purchase decision. Put another way, some consumers have an implied discount rate below 5%. In fact, a value of $m = 1.54$, is equivalent to having a negative implied discount rate under our assumptions. A negative value of m implies not only that (some) consumers do not value energy costs in their purchase decision and they choose energy-inefficient models over cheaper more energy-efficient models of similar quality. We provide evidence of existence of this strong degree of inattention in Section 5.

Table 3 reports the estimated shares of consumers that belong to four different types: consumers that are subject to behavioral biases such that $m < 0$, consumers subject to misperceptions that leads to an undervaluation of energy costs ($m \geq 0$, $m < 0.75$),[10] consumers that have modest misperceptions ($m \geq 0.75$, $m < 1.25$),[11] and consumers with misperceptions that leads to an overvaluation of energy costs ($m \geq 1.25$). The shares of these groups are 12.6%, 40.2%, 37.9%, and 18.9%, respectively. Therefore, although the mean level of misperception, $m = 0.77$ suggests a modest level of undervaluation, a significant share of consumers, > 60% with misperceptions that lead to under or over-valuation of energy costs.

The joint pdf of η and m (Panel a) shows that a large value of m is correlated with a lower coefficient on price. As we show below, this inverse correlation is partly, but not entirely, driven by the difference in the distribution of misperception across income groups. Several phenomena can explain overvaluation of energy costs. For instance, some consumers might pay much higher electricity prices to the ones that we have assumed in building the estimator, might have beliefs about current or future energy prices that are upward biased, may overestimate the energy consumption of refrigerators, or simply value energy efficiency highly because of environmental motives.

The marginal pdf of η (panel c) if left-skewed and has most of its mass between its mean value (-5.43) and zero. Some consumers have a really high price elasticity, some others are also price inelastic. We also find a positive and statistically significant mass for positive values of the price coefficient, which suggests that there is a share of consumers that are inattentive to the purchase

[10]Under our assumptions, a value of $m = 0.75$ corresponds to an implied discount rate of approximately 9%. Therefore, values of m between 0 and 0.75 translates into discount rates of 9% and above.

[11]This ranges for m translates into an implied discount rate between 3% to 9%.

price and/or have high search costs. Interestingly, consumers that have a negative value for m are also the ones that have a positive or a largely negative price coefficient. This suggests that these consumers are either focusing on dimensions of the product other than price or energy costs, or put a very strong emphasis on price alone, and dismiss energy costs completely.

4.3. **Robustness**

We investigate the robustness of the estimated pdfs by first considering alternative grid size and grid density. In Table 3, we report that result for estimators where we first reduced the span of the support and reduced the size of the grid (Model II).[12] We find that this increases the bunching at the lower end of the support, but it does not affect the mean and the overall patterns for the pdf distributions (joint and marginal) (Figure 6). In another specification (Model III), we kept the the same span for the support, but increased the number of grid points. With denser grid, we find more mass in the range $m \geq 0$, $M < 0.75$, but it has little impact on the mean (Table 3) and the qualitative patterns (Figure 6). Finally, we also consider a specification where the preference parameter for the ENERGY STAR label (τ) is also heterogeneous. Doing so allows us to further control for unobserved heterogeneity related to energy efficiency that might be correlated with preferences for energy costs. This has little effect on the estimated distributions of η and m (Figure 6)—we find similar patterns for the joint and marginal distributions of η and m.

4.4. **Distributions of Misperceptions by Income**

Table 3 summarizes the distribution of m across four regions of the parameter space that map into the four broad types of misperception for each of the six income groups. As discussed above, values of m smaller than zero indicate severe misperception. We find that this share of consumers is largest for the lowest income group, almost 30%. For other income groups, the estimated distribution suggests that this share ranges from 12% to 19%. Values of m between zero and 0.75, correspond to consumers that consider energy operating costs in their purchase decision, but undervalue them. The share of consumers in this category is larger among the lowest income groups and decreases

[12]We also attempted to implement estimator with a very large minimum value and maximum value for the support of the distribution, but we could not find a numerical solution for these estimators. The specification used for Model 1 in Table 3 is the largest span of the support for which we could recover estimates.

with income. For m between 0.75 and 1.25, which corresponds to no or modest misperception, the share of consumers is the largest among the three highest income groups. Finally, for m greater than 1.25, which implies overvaluation, the share of consumers increases with income and is especially large, higher than 30%, for income greater than $100k.

Figure 7 shows the joint distribution of η and m for each income group. The patterns are strikingly similar. The distributions are clustered around the mean, but there exist several smaller modes far from the means. We still observe that values of m that exceed one or are below zero are associated with lower coefficient on price. Consumers that respond strongly to energy costs are completely inattentive to energy costs tend to be also the less price sensitive (i.e., have a value of η close to zero). Again, for all income groups, we also observe a share of consumers that are very purchase price sensitive and those consumers tend to have m close to zero or negative. This type of consumer would favor cheaper appliances irrespective of their energy costs.

5. The Role of Inattention and Biased Beliefs

In this section, we explore to potential mechanisms behind the severe misperception that we observe: inattention and lack of information. We show that inattention may play a role, especially in explaining negative values of m. We also show that some consumer beliefs about key pieces of energy information are severely biased and follow patterns consistent with values of m that largely exceed one.

5.1. Inattention

In our discrete choice framework, a negative value for m implies that some consumers prefer energy-inefficient products over efficient ones, holding all other dimensions of quality constant. Therefore, in order to rationalize the probability mass of m over a negative support, our controls for the various dimensions of quality are crucial. In our estimator, the consumer-specific mean quality for attributes other than price and energy cost, γ_{ij}, are characterized by the product fixed effects and interaction terms between demographics and product attributes. The consumer-specific idiosyncratic quality is characterized by the error term that follows an extreme value distribution. Negative values of m arise when a consumer chooses an energy-inefficient product over other available options that have lower energy costs but higher quality, as suggested by their estimated mean quality, and given the dispersion of the idiosyncratic error term. Put simply, it appears that some consumers choose a dominated option both in terms of price and quality, and the only way the model can rationalize this decision is by having a negative value for m, which implies that energy costs were a desirable attribute.

In practice, inattention could be one of the underlying behavioral mechanisms that induces consumers to choose a less efficient product over a more efficient one of higher quality. For instance, if consumers dismiss energy costs in their purchase decision, and focus primarily on other dimensions of quality, this could lead to instances where they would make what appears to be a mistake by selecting a dominated option. It is also possible that consumers allocate attention to a restricted set of options, which results in having a dominant option excluded from the consideration set.

The empirical challenge in identifying these various types of inattention, and the resulting mistakes it induces, is that preferences for all dimensions quality need to be correctly specified. In our setting, the negative values for m in the FKRB estimator would accurately capture inattention only if our characterization of the mean quality, γ_{ij}, and the parametric assumption of the distribution of the error term provide a good approximation of the true preferences.

To investigate the role of inattention in our setting we propose an alternative empirical strategy that requires minimal assumptions about how we specify preferences for quality. We exploit a natural experiment in the U.S. refrigerator market, which allows us to compare choices among pairs of nearly identical refrigerator models, where a clear dominant and dominated options exist. In the U.S. refrigerator market, manufacturers commonly make strategic product line decisions to meet the Energy Star certification (Houde 2018a). This often results in product line with several different models, where some models only differ with respect to their energy consumption, but are otherwise identical from the point of view of the consumers. In our sample, we were able to identify 52 pairs of such refrigerators using a conservative matching process. To identify these pairs we first used the rich attribute information to ensure that paired refrigerators where identical along observable dimensions of quality, except energy use. In particular, we focus on product lines where manufacturers models with the same refrigerator volume, freezer volume, width, height, depth, overall design, color, and technology options but consume different amount of electricity. For all paired refrigerators identified by this matching process, we then manually verified the accuracy of the matches using information from various online marketplaces.

A second institutional detail that offers an ideal setting to detect inattention is the fact that the retailer has a national pricing strategy and offers large and frequent promotions. Therefore, each model is subject to large price variation and it is often the case that the price of the most efficient model within a pair is lower than the price of the less efficient model. When this occurs, this is a clear instance where consumers face a dominant and dominated option.

The existence of the identical pairs and model-specific price variation over time allows us to construct a fully non-parametric estimator that controls for all dimensions of quality so that we can estimate inattention to energy cost and price. To build such estimator, we first identify instances when the more efficient model was cheaper than the less efficient model for each pair. We refer to this type of event as a dominated price event. In then identify cases where both models within a pair were offered at the same location during a dominated price event. One challenge here is that we do not observe inventory and we must impute product availability from sales. However, the fact that we have transaction level data where we observe the exact location and date of each purchase allows us to address this issue as follows.

We infer that two models of a matched pair were available at the same location if we observe at least one sales for each model during a time period of specific length. We consider different lengths starting from the most conservative, where we only consider instance when both models sold on the same day at a given location, and we gradually increase the length to two, four, six, and twelve days. As we increase the time interval between sales, it becomes more likely that if the dominant product was not purchased, it is because of it was temporarily out-of-stock, rather than consumer inattention.

Once we have identified the dominate price events for each pair and restrict such events to locations where both models were offered at the same time, our non-parametric estimator of inattention simply consists of reporting the share of consumer that chose the dominated option.

Table A2 reports summary statistics for the matched pairs compared to the overall sample of models we observe. One take-away is that matched refrigerator models tend to be smaller and cheaper relative to the overall sample and, therefore, are not fully representative of the U.S. market. A second important take-away is that the average price difference within a pair is \$22 a small and positive amount, which is consistent with the fact that more efficient models tend to be more expensive. But there are large variations that span negative and positive values. The 10^{th} percentile is negative \$150 and the 90^{th} percentile is \$155.

Table 4 shows the share of consumers that chose the dominated option, along with the price difference and the number of observations during a dominated price event. For the most conservative assumption about product availability, where we only consider that at least one sales of both models occurred on the same day, we have only 304 observations that fulfill this criterion. This is a very small fraction of sales of the several million of transactions that we started with. Arguably, this estimator strives for internal validity and we err on the conservative side to truly identify inattentive consumers. We find that the share of inattentive consumers is 48%. As we increase the length of time between sales within a pair, the share of inattentive consumers tend to decrease and stabilize. It, however, remains large in magnitude. Across the various specifications, the share ranges from 48% to 38%.

These results show that consumers fail to systematically find about a clearly dominated option in their choice set, which highlights that inattention is at play in this setting. We, however, should be cautious in interpreting these results for the following reasons. First, the estimator captures inattention for particular options in the choice set, which could be induced by inattention to energy costs and prices, and by how consumers restrict their consideration set. Although, we do not distinguish between these two mechanisms, we argue that both are the result of imperfect attention allocation. Our estimator is thus a useful diagnostic to show that inattention is present and important in our setting. Second, although the estimator controls for all dimensions of quality, it does not control for the retailer's product placement strategies. In the absence of attention allocation costs though, product placement strategies should have a minimal impact. Such strategies can thus be interpreted as an equilibrium response to inattention. Our estimator is thus still relevant even if we do not identify how the retailer exacerbates or alleviates this behavioral bias. Third, the estimation is carried on a very small subset of the overall sample. The external validity of the results, therefore, is not guaranteed.

5.2. **Biased Beliefs**

Computing the energy cost of a refrigerator requires two key pieces of information: estimates of electricity consumption and electricity price. Beliefs about these pieces of information should then play an important role on how consumers respond to energy costs when they purchase an appliance.

In this section, we report data from a survey designed to assess energy literacy among U.S. households. The survey was administered in the Spring of 2017 on a representative sample of 1,512 U.S. households living in 24 different U.S. states. In this paper, we report the results for two questions that were included in the questionnaire and focused on beliefs. In the first question, we asked survey respondents to provide their best estimate of a full-size refrigerator's annual electricity use. In the second question, we asked them for their best estimate of the average electricity price that they pay. For both questions, respondents were asked to report their best estimates and no numerical or other anchors were provided.

Figure 8 reports the distribution of the ratio between beliefs and true value for the two pieces of information. Allcott (2013) refer to such measures as valuation ratios, where correct valuation implies a ratio of exactly one. To construct our valuation ratios, we matched county average electricity prices using information about respondents' zip codes. For refrigerator's annual electricity use, we simply use the value 450 kWh/year, which approximately corresponds to average for the year 2017. For both pieces of information, we see a large distribution of beliefs, with significant under- and over-valuation. To construct the histograms, we censored all values greater than 5, which explains the large mass at this value. This shows that a significant share of consumers tend to have beliefs about refrigerators' energy use and local electricity prices that would lead to an overvaluation of energy costs. The pattern in beliefs is u-shaped and also suggests that a large share of consumers undervalue both the level of energy use of refrigerator and prices. Interestingly, Allcott and Knittel (2018) find similar patterns in beliefs in the U.S. car market.

In sum, we find that beliefs are severely biased in our setting and follow patterns consistent with the distribution of misperception our estimator recovers. In particular, the large values of m could be induced by consumers that overestimate energy use and electricity price. Biased beliefs are, however, one possible mechanism. As discussed earlier, environmental values could also play a role.

It is also important to note that in the U.S. appliance market, the mandatory label EnergyGuide should help informing consumers about different pieces of energy information. How effective is the label in correcting beliefs remain an open question, but our survey results show that there is a need to provide better information.

6. Implications for Policy Design

In this section, we assess how heterogeneity in misperception of energy costs impacts the design of policies used to address negative externalities associated with energy use. To illustrate the role that heterogeneity plays, we focus on simple cases where the planner uses a single policy instrument, a Pigouvian tax versus different type of standards. We show how heterogeneous misperceptions affect the level of the optimal instruments and their ranking.

Our measure of welfare is a direct application of Leggett (2002)'s formula to measure welfare in a discrete choice framework in the presence of imperfect information. This framework has been further developed by Allcott (2013), Ketcham et al. (2016b), and Houde (2018b) to measure welfare in the presence of consumers' biases. The first step in applying this framework is to make a distinction between decision and experienced utility. Decision utility refers to what consumers thought they will experience upon making a purchase decision. Experienced utility refers to the true flow of utility that consumers will experience from their purchase, irrespective of their beliefs at the time they made a purchase decision. The welfare effect of a policy should then be based on experienced utility. If consumers misperceive some components of product costs, decision and experienced utility will differ. In this case, Leggett (2002) has shown that the change in consumer surplus for a given policy change can still be measured and expressed as the gap between decision and experienced utility:

$$\begin{aligned} \Delta CS_{ktr} = \frac{1}{\eta_k} \cdot \left[ln \sum_j^J exp(\tilde{U}_{kjtr}) + \sum_j^J \tilde{P}_{kjtr} \cdot (\tilde{U}^E_{kjtr} - \tilde{U}_{kjtr}) \right] - \\ \frac{1}{\eta_k} \cdot \left[ln \sum_j^J exp(U_{kjtr}) + \sum_j^J P_{kjtr} \cdot (U^E_{kjtr} - U_{kjtr}) \right]. \end{aligned} \tag{6}$$

where the terms with a tilde are evaluated after the policy change, U^E_{kjtr} denotes experienced utility, U_{kjtr} corresponds to decision utility for consumer of type k and P_{kjrt} refers to the probability that consumer type k will choose product j in region r in time t. The expression in 6 corresponds to the standard measure of welfare for the multinomial logit (Small and Rosen 1981) plus the term $\sum_j^J P_{kjtr} \cdot (U^E_{kjtr} - U_{kjtr})$, which we refer to as the correction term. The correction term arises because of the discrepancy between what consumers perceive they will experience and what they actually experience, and simply represents the expected (private) cost that consumers incur because of their misperceptions.

The second step in applying this framework is to take a stand on what we believe is an accurate representation of experienced utility. Because the estimation yields parameter estimates that characterizes decision utility, we must make assumptions to recover experienced utility. For our application, we will make the following two assumptions:

(1) If $m_k \neq 1$, consumer of type k misperceived energy costs and his decision utility differs from experience utility. To measure experience utility for type k, we set $m_k = 1$.

(2) If $\eta_k > 0$, consumer of type k misperceived the product price and decision utility differs from experience utility. To measure experience utility for type k, we set $\eta_k = \bar{\eta}_{-k}$, where $\bar{\eta}_{-k}$ refers to the average value for the coefficient η for all types other than k.

The first assumption implies that if consumers do not perceive a one dollar change in future energy operating costs (discounted with a normal rate of return) the same way they perceive a dollar change in product price, they are prone to a bias. Note that by setting $m_k = 1$, we still let consumers to be heterogeneous with respect to their response to price (i.e. marginal utility of income). The second assumption aims to deal with consumer types for which η is greater than zero. A positive coefficient on price is a manifestation of various consumers' biases with respect to the product price and notably high search costs. By setting $\eta_k = \bar{\eta}_{-k}$, we make the implicit assumption for all types $j \neq k$ for which $\eta_j \leq 0$, that the coefficient on price reflects their true marginal utility of income. Furthermore, we also implicitly assume that the distribution of the true marginal utility of income for consumers $\eta_k > 0$ follows the same distribution as the remaining of the population.[13]

[13]Ketcham et al. (2016b) propose an alternative approach to deal with consumers' biases pertaining to product price. They first devise a procedure to identify choices that are considered a mistake in their

6.1. **Optimal Pigouvian Tax**

Heterogeneity in misperceptions of energy costs has also implications for the design of policies that aim to correct negative externalities associated with energy use. For instance, Allcott et al. (2014) show that the optimal Pigouvian energy tax in the presence of biased perception of energy costs should be adjusted using information about the distribution of biases where the average bias is not a sufficient statistic to adjust the level of the tax. Houde and Aldy (2017) illustrate this result by deriving an expression for the optimal price-inclusive Pigouvian energy tax in the presence of misperceptions of energy costs. Using the expression 6 to measure consumer welfare, and assuming that tax revenues are redistributed lump-sum, the tax is included in the marginal price of energy, denoted P_e, and each unit of energy consumed (E_j) is associated with a constant marginal damage cost, denoted ϕ, the optimal Pigouvian tax, τ^*, is:

$$\tau^* = \frac{\phi}{1-\mathcal{A}} + P_e \frac{\mathcal{A}}{1-\mathcal{A}} \tag{7}$$

with

$$\mathcal{A} = \sum_k \alpha_k (1 - m_k) \sum_j \frac{\partial P_j^k}{\partial \tau} E_j$$

If instead the tax is set using the average degree of misperceptions, noted $\bar{m}$, the optimal (price-inclusive) Pigouvian energy tax[14] is simply:

$$\tau^* = \frac{\phi}{\bar{m}} + P_e \frac{1-\bar{m}}{\bar{m}} \tag{8}$$

Clearly, the expressions in 7 and 8 are different. In Appendix A4, we further show the setting a tax that solely relies on the average misperception, $\bar{m}$, and does not consider the distribution of misperceptions leads to a bias that is proportional to the the variance of m. Heterogeneity in

sample. They tend exclude these choices from the sample, and estimate a choice model on what they consider a sample free of consumers' mistakes. They treat the coefficient on price from this estimation as the true marginal utility of income.

[14]The expression 8 is similar to the results of Farhi and Gabaix (2018) except that the price of energy also enters the expression of the optimal Pigouvian tax. It can also be shown that Proposition 1 of Allcott et al. (2014) yields a similar expression to Equation 7, except that their expression for the optimal tax is function of the degree of misperception in future utilization of the durable, in addition of the price of energy, and misperception m_k.

misperception of energy costs has also implications for the choice of Pigouvian instruments for addressing negative externalities. For instance, it may be optimal to combine a tax with a subsidy (Allcott et al. 2014), or to favor a quantity instrument such as minimum standard over pricing the negative externalities (Allcott and Knittel 2018; Farhi and Gabaix 2018).

6.2. **Optimal Standard**

In our policy simulations, we consider three types of standards. First, we consider an uniform minimum standard where all products must consume the same amount of energy. This standard represents an extreme case where manufacturers cannot differentiate their products in the energy dimension.[15] Second, we consider a minimum standard where all products must meet a lower bound with respect to energy-efficiency, but can offer also products that exceed this threshold. Alternatively, this standard dictates the maximum amount of energy a given product can consume, but does not restrict the minimum consumption. Third, we consider a variant of the minimum standard where the standard varies as a function of key attributes. This attribute-based standard mimics the regulations currently in place in the U.S. appliance market.

Under our framework, an expression for the optimal uniform standard can be easily derived. We will focus on showing this result, which provides much of the intuition on the role of heterogeneous misperceptions in deciding what should be the optimal stringency and design of a standard. To derive our main result, we will assume a very stylized market structure and abstract from the strategic behavior of the firms. Our goal is to focus on the role of the demand side. We assume that manufacturers set price such that the purchase price of product j is given by:

$$p_j(E_j) = c(E_j) + \omega_j \tag{9}$$

where $c(e)$ is the product cost that varies as a function of the energy level E_j, with $c'(E) < 0$, and ω_j is a product-specific additive markup that does not vary with E_j. As before, we will assume a linear and additive externality cost and that the existence of k different types subject to bias m_k.

[15]The government of India has implemented a version of such standard by requiring that only one variant of CFL light bulb be offered. They are currently considering a similar approach for energy-efficient air-conditioners where only one variant would be offered.

Each type k represents a share α_k of the overall population. The following proposition shows the expression for the optimal uniform standard under this setup.

Proposition 2. *When $p_j(E_j) = c(E_j) + \omega_j$ and with a constant marginal externality cost, ϕ, the optimal uniform standard, denoted $\bar{E}^*$ is:*

$$\bar{E}^* = \frac{c'(E)}{1+\phi}. \tag{10}$$

The important takeaway from Equation 10 is that the optimal standard is not function of the degree of misperceptions. That is, this result holds irrespective of the distribution of m_k. The design of an optimal uniform standard is thus unaffected by misperceptions.

This is because, by setting a uniform standard, all options have the same energy level—therefore, energy operating costs do not induce substitution across products. As a result, misperceptions of energy costs are completely internalize no matter their distribution. The only factors that matter in setting the optimal standard is the trade-off between the increase in product cost induced by making the standard more stringent (i.e., $c'(E)$) and the externality cost ϕ. In comparison to a price instrument, a standard has then a clear advantage given its ability to internalize misperceptions, even if there are heterogeneous, while simultaneously addressing the negative externalities.

The above result is specific to a very restrictive type of standard, but provides important insights on how other types of standards will be affected by heterogeneity in misperceptions. As standards reduce the variance in energy costs between products present in the choice set, this reduces the distortionary effects from misperception in trading off energy costs with other attributes. The uniform standard is the extreme case where there is a no variance in energy cost and all the substitution is ruled out. A minimum standard should induce some variation in energy operating costs, but less than a similar but attribute-based minimum standard. Therefore, the former should be more robust to the distribution of misperceptions relative to the latter.

Note that although a uniform standard rules out substitution across products, they may affect demand via other margins, namely the decision to adopt or not a product and its timing. A stringent uniform standard will raise the average purchase prices of a given technology, which could cause some consumers to reconsider adopting such technology. In markets, where those margins are important, it is no more true that misperceptions would not affect the level of the optimal standard.

6.3. **Setup: Policy Simulations**

We simulate the demand model to find the optimal level of the tax or different types of standards. For each policy, we compare two scenarios: the level of the optimal policy using the full distribution of misperceptions versus the average degree of misperceptions. For each scenario, we, however, always consider heterogeneity across income groups. The overall demand is thus a weighted average of the demand model estimated for each income group. Note all also that we consider heterogeneity in the share of each income group across all the U.S. counties. The demand model has spatial heterogeneity in the degree of misperceptions to energy costs due to difference in income across regions.

In addition to compare the level of the optimal policy, we also report total social welfare, consumer surplus, and externality costs, but we abstract from firms' profits. Across scenarios, we assume that firms' markups are fixed, and prices are set by: $p_j(E_j) = c(E_j) + \omega_j$. For simulating the standards, we need to model how energy consumption affects the products prices. For this purpose, we use the cost function estimated in Houde and Aldy (2017a), which takes the following parametric form:

$$c(E_j) = \frac{\psi}{E_j} + \beta_j \tag{11}$$

For all scenarios, we assume that there is a carbon externality and we fix the level of the externality at 50 $ per ton of CO2. We consider heterogeneous local electricity prices and emissions factors, which vary with electricity market areas. Data for the CO2 emission factors comes from the EIA scenarios used to simulate the impact of the Clean Power Plan on electricity prices. We infer the emission factors for 20 different regions of the U.S. corresponding to different electricity market. For electricity prices, we use the county averages constructed for our estimation.

To assess the robustness of our results with respect to our assumptions about experience utility, we consider an alternative scenario where we allow for heterogeneous discount rates. Our goal for this scenario is to be as conservative as possible regarding the possible existence of misperceptions. We consider a range of discount rate that spans 2% to 12%. For all values of m_k that correspond to implicit discount rates that fit this range, we assume that there is not misperception and then set $m_k = 1$. For values outside this range, we then use 2% as the basis to quantify overvaluation and 12% to quantify undervaluation.

Finally, we account for uncertainty due the estimation by sampling the different demand estimates obtained from the subsampling bootstrap procedure. We sample one set of demand estimates for each income group, simulate the optimal policy, store the results, and repeat. We report the mean and standard errors across a large number of repetitions (N=500).[16]

6.4. **Results**

Table 5 presents the results. For the tax instrument, the benchmark for the level of the optimal Pigouvian tax without misperceptions is 50 $/ton. If we consider the full distribution of heterogeneity in misperceptions, we found that the optimal tax is large and negative: -91.7 $/ton, and -91.4 $/ton if we consider heterogeneous discount rates. A negative tax means that we should subsidize the carbon externality in this market. This surprising result is driven by the fact that there is a large enough share of consumers that overvalue energy costs, which induces a large downward adjustment to the tax. If instead we ignore heterogeneity, and we set the tax based on the average value of m, the optimal tax ranges from 113.1 to 148.1 $/ton, which is more than two times the externality cost. Accounting for heterogeneity in misperceptions has thus a very dramatic effect on the level of the optimal tax. Note also that using different assumption to determine the level of experienced utility has very little effect, if we consider the whole distribution. The results are, however, much more sensitive if we only focus on the mean. Again, this is a cautionary tale about relying on a single statistic to determine the level of the policy.

[16]Note that for each income group, we have used 16 different slices of the data to implement the subsampling bootstrap. Given that there are 6 income groups, there are then 16^6 different combinations of the demand models that we can sample from for the policy simulations.

For standards, accounting for heterogeneity have very little effect on the results. For the optimal uniform standard, as shown in Proposition 2, the level is unaffected by heterogeneity in misperceptions and is set at 335 kWh/year. Compared to the average energy consumption observed in the sample, 514.9 kWh/year (Table 1), this corresponds to a reduction of about 35%. For the minimum standard, we found that optimal standard is fact the same than the uniform standard and again heterogeneity in misperceptions has no effect on the standard. This is not a theoretical results, but a special case for our setting. The fact that the minimum standard is very stringent implies that it is binding for all products. The minimum standard thus acts as an uniform standard. For the case of the attribute-based minimum standard, which we express as a percentage reduction relative to the existing minimum standard, the optimal standard ranges from 54% to 56% of the existing standard. We find that heterogeneous misperceptions have a small but negligeable effect on the level of standard.

Overall, all three types of standards are very robust to misperceptions and heterogeneity. Moreover, they induce welfare gains that are much larger than the Pigouvian taxes. One caveat of the above results is that we have abstracted from firms' profits. Standards impose costs directly to manufacturers and thus should have a negative impacts on profits, relative to a tax. We thus consider the results on Table 5 as an upper bound of the relative benefits of standards relative to a tax. Nonetheless, focusing on demand solely provide an important take-away. The ability of a quantity instrument to address misperceptions by reducing the variance in energy costs while simultaneously addressing the externality provides a clear advantage for this type of instrument.

7. Conclusion

The standard test of consumer misperception used widely in the literature compares the responsiveness of demand for changes in potentially misperceived aspects of cost against salient, correctly perceived aspects of cost. Consumers should be indifferent between an additional dollar of purchase price and an additional dollar of the potentially misperceived cost such as shipping and handling or, the present discounted dollar of energy expenditure, since total lifetime cost should be the relevant metric. The ratio of the responsiveness coefficients has been used as a sufficient statistic for the average degree of consumer misperception.

This paper shows that the average degree of misperception can be a misleading statistics. Empirically, it can hinder large misperceptions. From a policy design standpoint, ignoring heterogeneity and focusing on the average, can lead also to very different conclusions. We illustrate these points using the U.S. appliance market and perception of energy operating costs as a case study. We show that the valuation of energy operating costs are very heterogeneous and suggest large misperceptions ranging from undervaluation to large overvaluation. The average degree of misperceptions, however, suggests a very modest amount of misperceptions.

We use these results to show how different policies that aim to address the carbon externality associated with energy use while accounting for misperceptions are affected by heterogeneity. We show heterogeneous misperceptions have a large impact on the level of the Pigouvian tax, but very little effect on a quantity instrument such as a standard. In fact, we show that a special type of standard, uniform standard, could even internalize misperceptions and externalities all at once. Empirically, we find that different types of standards performs relatively well from a welfare standpoint and dominate a tax. Ignoring firms' profits, we also find support for much stringent energy efficiency standards.

References

Abaluck, Jason and Jonathan Gruber, "Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program," *American Economic Review*, 2011, *101* (4), 1180–1210.

Allcott, H., "The Welfare 3ffects of Misperceived Product Costs: Data and Calibrations from the Automobile Market," *American Economic Journal: Economic Policy*, 2013, *5* (3), 30–66.

———— **and C. Knittel**, "Are Consumers Poorly Informed about Fuel Economy? Evidence from Two Experiments," *American Economic Journal: Economic Policy*, 2018, *forthcoming.*

Allcott, Hunt and Dmitry Taubinsky, "Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market," *American Economic Review*, 2015, *105* (8), 2501–2538.

———— **and Nathan Wozny**, "Gasoline Prices, Fuel Economy, and the Energy Paradox," *The Review of Economics and Statistics*, 2014, *96* (5), 779–795.

————**, Sendhil Mullainathan, and Dmitry Taubinsky**, "Energy Policy with Externalities and Internalities," *Journal of Public Economics*, 2014, *112* (April), 72–88.

Allen, Jason, Robert Clark, and Jean-François Houde, "Search frictions and market power in negotiated price markets," Technical Report, National Bureau of Economic Research 2014.

Barber, Brad M., Terrance Odean, and Lu Zheng, "Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows," *Journal of Business*, 2005, *78* (6), 2095–2120.

Bernheim, B. Douglas and Dmitry Taubinsky, "Behavioral Public Economics," *NBER Working Paper*, 2018, (24828).

Blundell, Wesley, Gautam Gowrisankaran, and Ashley Langer, "Escalation of Scrutiny: The Gains from Dynamic Enforcement of Environmental Regulations," Technical Report, National Bureau of Economic Research 2018.

Bollinger, Brian, Phillip Leslie, and Alan Sorensen, "Calorie Posting in Chain Restaurants," *American Economic Journal: Economic Policy*, 2011, *3* (1), 91–128.

Busse, Meghan, Christopher Knittel, and Florian Zettelmeyer, "Are Consumers Myopic? Evidence from New and Used Car Purchases," *American Economic Review*, 2013, *103* (1), 220–256.

Chetty, R., A. Looney, and K. Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 2009, *99* (4), 1145–1177.

Datta, S. and M. Filippini, "Analysing the Impact of ENERGY STAR Rebate Policies in the U.S.," *Energy Efficiency*, 2016, *9*, 677–698.

DellaVigna, Stefano, "Structural behavioral economics," Technical Report, National Bureau of Economic Research 2018.

Ellison, Glenn, "A Model of Add-on Pricing," *Quarterly Journal of Economics*, 2005, *120* (2), 585–637.

——— **and Sara Fisher Ellison**, "Search, Obfuscation, and Price Elasticities on the Internet," *Econometrica*, 2009, *77* (2), 427–452.

Farhi, E. and X. Gabaix, "Optimal Taxation with Behavioral Agents," *NBER Working Paper*, 2018, (21524).

Gabaix, Xavier and David Laibson, "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, 2006, *121* (2), 505–540.

Gertler, Paul J., Orie Shelef, Catherine D. Wolfram, and Alan Fuchs, "The energy-efficiency gap What does it mean?," *American Economic Review*, 2016, *106* (6), 1366–1401.

Grigolon, Laura, Mathias Reynaert, and Frank Verboven, "Consumer Valuation of Fuel Costs and the Effectiveness of Tax Policy: Evidence from the European Car Market," *American Economic Journal: Economic Policy: Forthcoming*, 2018.

Guiso, Luigi, Andrea Pozzi, Anton Tsoy, Leonardo Gambacorta, and Paolo Mistrulli, "The Cost of Steering in Financial Markets: Evidence from the Mortgage Market," 2018.

Handel, Benjamin R. and Jonathan T. Kolstad, "Health Insurance for "Humans":Information Frictions, Plan Choice, and Consumer Welfare," *American Economic Review*, 2015, *105* (8), 2449–2500.

Harrison, Glenn W and John A List, "Field experiments," *Journal of Economic literature*, 2004, *42* (4), 1009–1055.

Heiss, Florian, Daniel McFadden, Joachim Winter, Amelie Wuppermann, and Bo Zhou, "Inattention and Switching Costs as Sources of Inertia in Medicare Part D," Working Paper 22765, National Bureau of Economic Research October 2016.

Hossain, T. and J. Morgan, "Plus Shipping and Handling: Revenue (Non)Equivalence in Field

Experiments on eBay," *Advances in Economic Analysis and Policy*, 2006, *6* (2).

Houde, S. and J.E. Aldy, "The Efficiency Consequences of Heterogeneous Behavioral Responses to Energy Fiscal Policies," Working Paper 24103, National Bureau of Economic Research December 2017.

Houde, Sébastien, "Bunching with the Stars: How Firms Respond to Environmental Certification," *Economics Working Paper Series*, 2018, *18* (292).

Houde, Sébastien, "How Do Consumers Respond to Product Certifications and Value Energy Information," *RAND Journal of Economics*, 2018, *49* (2), 453–477.

——— **and Erica Myers**, "Are Consumers Attentive to Operating Costs? Evidence from the Appliance Market," *NBER Working paper*, 2018.

Houde, Sebastien and Joseph E. Aldy, "Consumers' Response to State Energy Efficient Appliance Rebate Programs," *American Economic Journal: Economic Policy*, 2017, *9* (4), 227–255.

Jensen, Robert, "The Perceived Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics*, 2010, *125* (2), 515–548.

Ketcham, Jonathan D, Nicolai V Kuminoff, and Christopher A Powers, "Estimating the Heterogeneous Welfare Effects of Choice Architecture: An Application to the Medicare Prescription Drug Insurance Market," Working Paper 22732, National Bureau of Economic Research October 2016.

Ketcham, Jonathan D., Nicolai V. Kuminoff, and Christopher A. Powers, "Estimating the Heterogeneous Welfare Effects of Choice Architecture: an Application to the Medicare Prescription Drug Insurance Market," *NBER Working Paper*, 2016, (22732).

Kling, Jeffrey, Sendhil Mullainathan, Eldar Shafir, Lee Vermeulen, and Marian Wrobel, "Comparison Friction: Experimental Evidence from Medicare Drug Plans," *Quarterly Journal of Economics*, 2012, *127* (1), 199–235.

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

Myers, E., "Are Home Buyers Inattentive? Evidence From Capitalization of Energy Costs," *American Economic Journal: Economic Policy: Forthcoming*, 2018.

Nevo, Aviv, John L Turner, and Jonathan W Williams, "Usage-Based Pricing and Demand for Residential Broadband," *Econometrica*, 2016, *84* (2), 411–443.

Revelt, D. and K. Train, "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level," *Review of Economics and Statistics*, 1998, *8* (4), 647–657.

Sallee, J.M., S. West, and W. Fan, "Do Consumers Recognize the Value of Fuel Economy? Evidence from Used Car Prices and Gasoline Price Fluctuations," *Journal of Public Economics*, 2016, *135*, 61–73.

Small, K.A. and H.S. Rosen, "Applied Welfare Economics with Discrete Choice Models," *Econometrica*, 1981, *49* (1), 105–130.

Taubinsky, D. and A. Rees-Jones, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," *Review of Economic Studies: Forthcoming*, 2018.

Train, Kenneth E, *Discrete choice methods with simulation*, Cambridge university press, 2009.

8. Tables and Figures

Table 1. Summary Statistics

	Mean	SD
Price ($)	1252.6	627.0
kWh/y	514.9	78.4
County Elec. Price (cents)	11.4	3.7
State Elec. Price (cents)	12.3	3.3
County Elec. Cost/y ($)	58.5	20.6
State Elec. Cost/y ($)	63.2	18.9
Rebate Amount ($)	25.9	68.8
% Energy Star	68.5	
% w Ice-Maker	76.0	
Overal Size (cu. ft.)	22.5	3.4
% w Top Freezer	30.3	
Demographics		
% of Households	67.6	
% w. Demo. Info.	56.6	
% Renters	1.9	
Income distribution		
<$30k	12.2	
$30k-$50k	16.8	
$50k-$75k	25.2	
$75k-$100k	18.2	
$100k-$150k	11.8	
>$150k)	15.7	

Table 2. Poisson and Conditional Logit Regressions

	CLogit-I	CLogit-II	CLogit-III	CLogit-IV
Purchase Price	-0.00348***	-0.00348***	-0.00348***	running
	(0.00007)	(0.00007)	(0.00007)	
Energy Cost	-0.0290***	-0.0305***	-0.0309***	running
	(0.00318)	(0.00318)	(0.00318)	
Misperception: m	0.7131	0.7485	0.7591	
Product FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Brand × Week FE	No	No	No	Yes
County × ES FE	No	No	Yes	Yes
Demo × EE-related Attributes	No	Yes	Yes	Yes

Notes: Standard errors (in parentheses) clustered at the county level. The average misperception is computing assuming a discount rate of 5% and a refrigerator lifetime of 18 years.

Table 3. Marginal PDF of m: Robustness and Heterogeneity w.r.t. Income

	m<0	m≥0 m<0.75	m≥0.75 m<1.25	m≥1.25	E[m]
Model I	12.63	40.19	37.92	18.88	0.77
	(2.27)	(5.84)	(6.80)	(3.57)	(0.08)
Robustness Tests					
Model II: Smaller Grid	12.94	39.63	43.50	20.25	0.73
	(2.20)	(5.52)	(5.87)	(3.10)	(0.07)
Model III: Dense Grid	13.56	48.00	31.91	19.63	0.76
	(2.15)	(5.68)	(8.31)	(3.53)	(0.08)
Model IV: τ Heterogeneous	15.83	37.00	33.78	21.83	0.78
	(2.67)	(7.85)	(8.08)	(5.34)	(0.10)
Heterogeneity w.r.t. Income					
< \$30k, Model I	27.43	56.43	14.00	7.17	0.35
	(6.71)	(10.64)	(8.13)	(2.56)	(0.07)
\$30-50k, Model I	17.81	64.69	17.43	11.14	0.48
	(3.82)	(5.45)	(6.29)	(3.05)	(0.07)
\$50-75k, Model I	12.75	53.31	25.11	21.00	0.73
	(2.70)	(6.33)	(6.80)	(3.14)	(0.10)
\$75-100k, Model I	15.31	44.47	32.30	23.00	0.82
	(2.05)	(5.00)	(7.16)	(3.17)	(0.07)
\$100-150k, Model I	19.56	31.15	27.00	36.81	0.94
	(4.61)	(6.71)	(6.40)	(6.22)	(0.12)
≥ \$150k, Model I	13.38	43.71	31.13	32.88	1.00
	(2.48)	(5.55)	(7.75)	(6.96)	(0.12)

Notes: Marginal PDF of m computed from the joint PDF of η and m. The two-step estimation was performed for model and for each income group separately. Standard errors (in parentheses) are obtained using subsampling bootstrap. The first four columns refer to different bins of the pdf, which each maps into a different type of misperception. The last column is the average amount of misperception implied by the full distribution.

Table 4. Matched Pairs: Share of Inattentive Consumers

	Time Interval Between Sales Within Pair						
	Same Day	+/- 1 Day	+/- 2 Days	+/- 4 Days	+/- 6 Days	+/- 12 Days	+/- 24 Days
Dominated Option	48.03%	45.85%	44.44%	43.31%	43.18%	41.27%	38.27%
Δ Price (\$)	-19.70	-21.86	-25.05	-29.38	-33.13	-39.24	-45.61
# Obs	304	325	360	441	528	773	1304

Table 5. Optimal Behavioral Policies

	Misperceptions at $r = 5\%$				Misperceptions with Heterogeneous r			
	Policy	Δ SW $/capita	Δ CS $/capita	Δ Ext $/capita	Policy	Δ SW $/capita	Δ CS $/capita	Δ Ext $/capita
Pigou Tax ($/ton of CO_2)								
With $F(m_k)$	-91.7	3.5	209.4	1.3	-91.4	3.9	208.7	1.4
	(0.38)	(0.05)	(0.90)	(0.01)	(0.38)	(0.06)	(0.89)	(0.01)
With $E[m_k]$	113.1	2.6	-237.7	-2.2	148.1	3.9	-312.4	-2.6
	(2.05)	(0.07)	(4.27)	(0.03)	(1.80)	(0.08)	(3.76)	(0.03)
Uniform Standard (kWh/year)								
With $F(m_k)$	335.3	128.3	91.0	-37.3	335.3	128.0	90.8	-37.1
	(0.00)	(0.13)	(0.13)	(0.01)	(0.00)	(0.15)	(0.14)	(0.02)
With $E[m_k]$	335.3	100.8	65.9	-34.9	335.3	103.2	67.4	-35.8
	(0.00)	(0.10)	(0.09)	(0.03)	(0.00)	(0.11)	(0.09)	(0.03)
Minimum Standard (kWh/year)								
With $F(m_k)$	335.3	128.3	91.0	-37.3	335.3	128.0	90.8	-37.1
	(0.00)	(0.13)	(0.13)	(0.01)	(0.00)	(0.15)	(0.14)	(0.02)
With $E[m_k]$	335.3	100.8	65.9	-34.9	335.3	103.2	67.4	-35.8
	(0.00)	(0.10)	(0.09)	(0.03)	(0.00)	(0.11)	(0.09)	(0.03)
Attribute-Based Minimum Standard (% w.r.t. existing standard)								
With $F(m_k)$	54.7	98.2	61.9	-36.3	55.0	99.5	63.6	-35.8
	(0.03)	(0.11)	(0.09)	(0.04)	(0.03)	(0.10)	(0.08)	(0.04)
With $E[m_k]$	56.1	79.7	47.1	-32.6	56.1	80.9	47.8	-33.1
	(0.01)	(0.11)	(0.10)	(0.02)	(0.01)	(0.10)	(0.10)	(0.02)

Notes: The table reports the optimal policies evaluated using the full distribution of misperceptions (with $F(m_k)$) or only the average misperception (with $E[m]$). For each of the two cases, the change in social welfare (SW), consumer surplus without redistribution of the tax revenues (CS), and externality costs (Ext) are computed relative to the case without a policy. The optimal taxes, standards, and welfare metrics are computed for random draw of the estimated distributions recovered by the subsambpling bootstrap. The mean and standard errors (in parentheses) across 500 draws are reported.

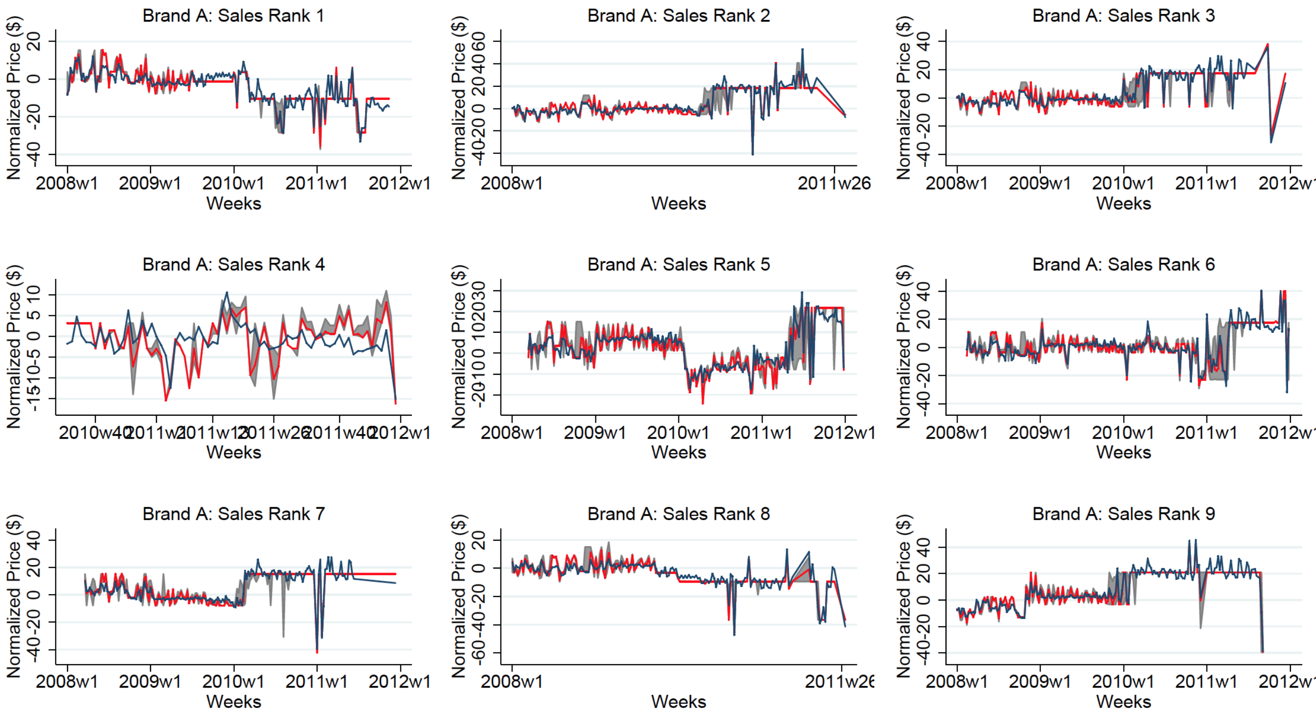


FIGURE 1. Price Variation

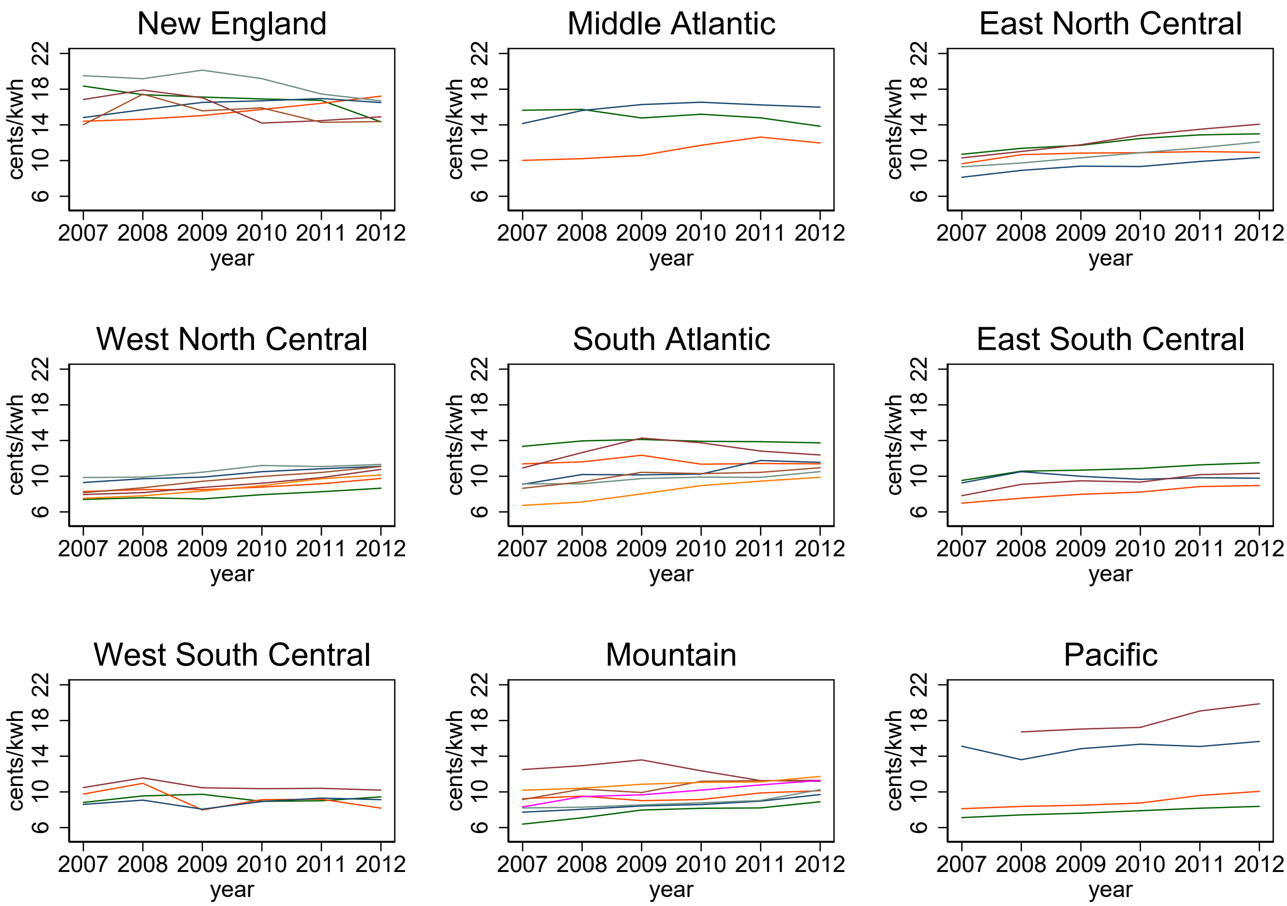


FIGURE 2. Average Electricity Prices for Each State in a Census Divsion

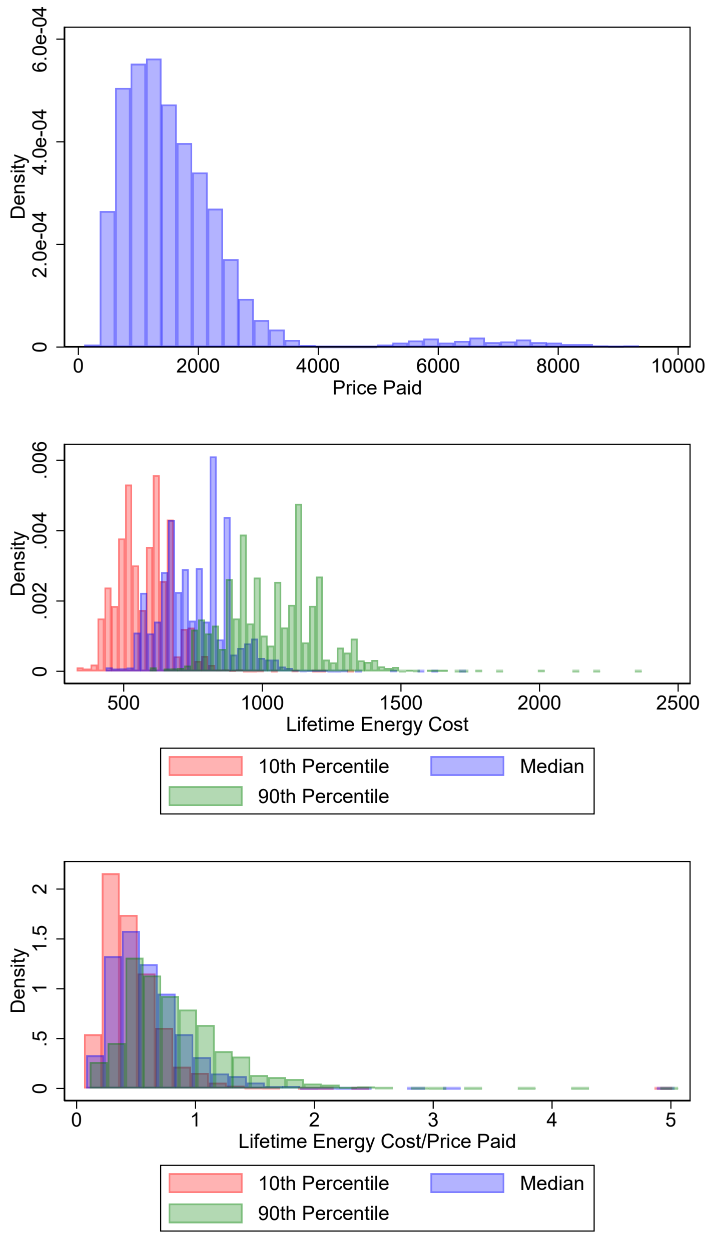


Figure 3. Distributions of Prices and Lifetime Energy Costs

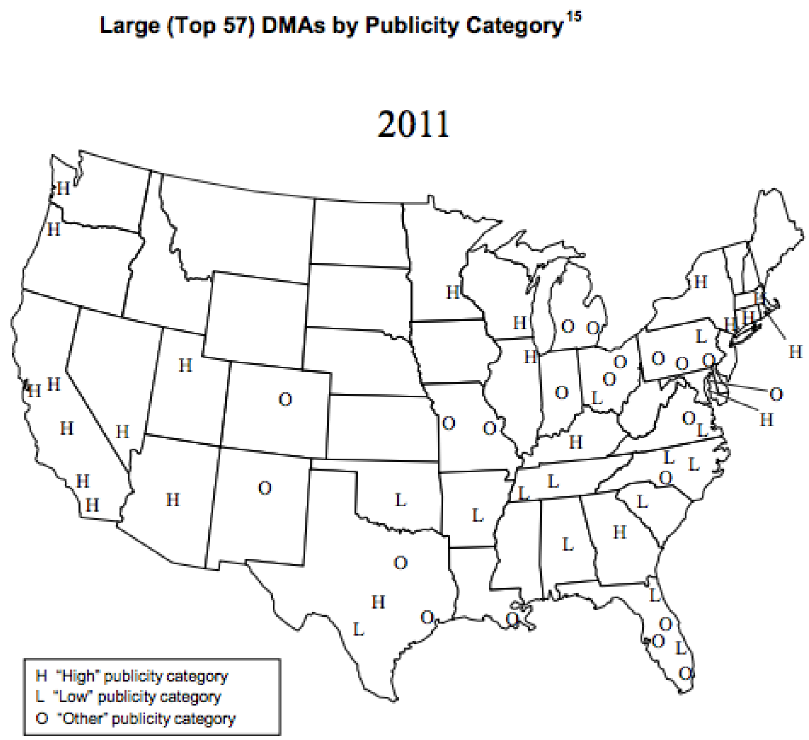
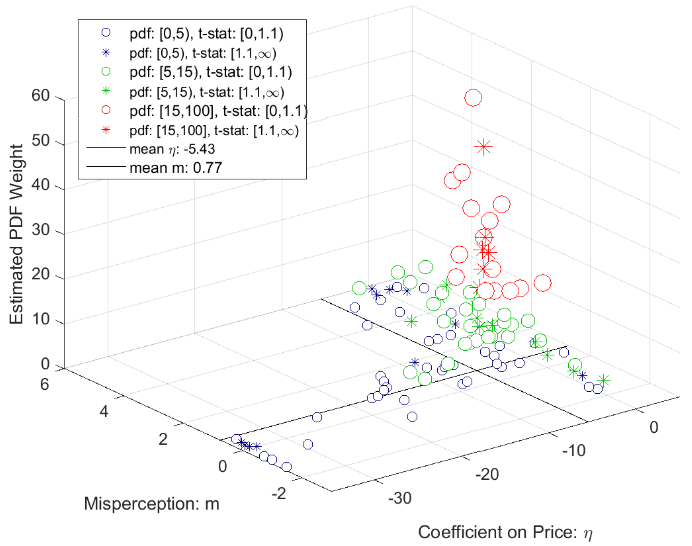
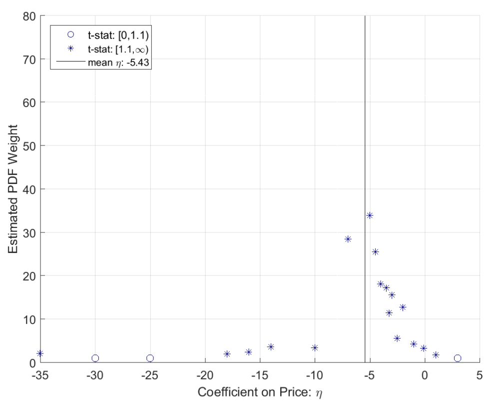


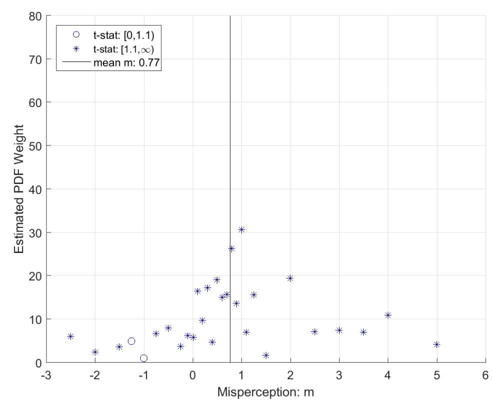
Figure 4. Source: EPA



(a) Joint PDF of η and m: $f(\eta, m)$



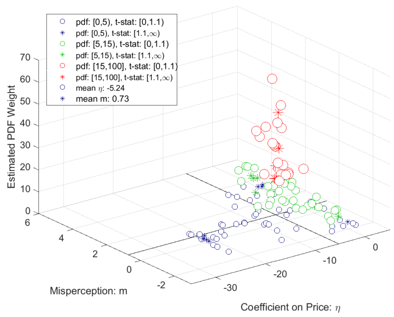
(b) Marginal PDF of η: $f(\eta)$



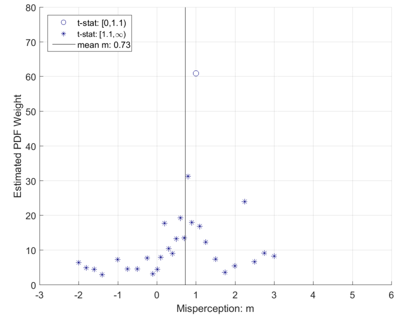
(c) Marginal PDF of m: $f(m)$

FIGURE 5. Joint and Marginal Probability Distribution for η and m

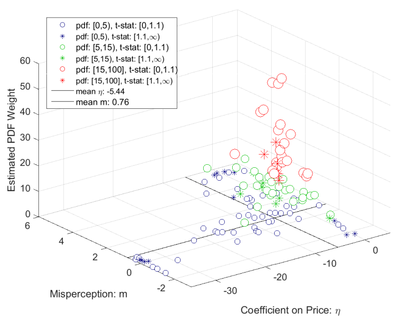
Notes: : .



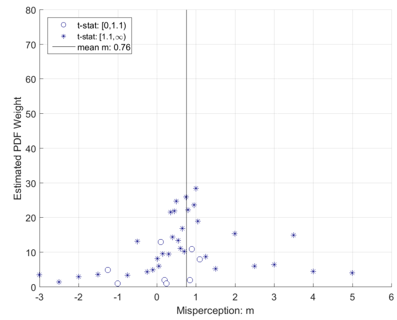
(a) Model II Smaller Span; $f(\eta, m)$



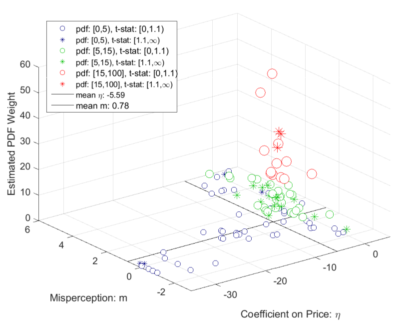
(b) Model II Smaller Span; $f(m)$



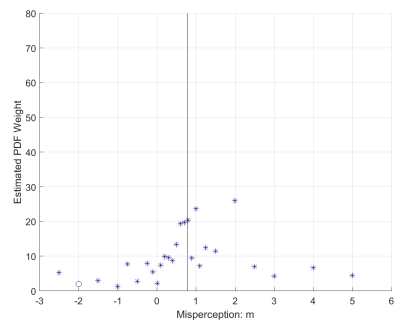
(c) Model III Dense Grid; $f(\eta, m)$



(d) Model III Dense Grid; $f(m)$

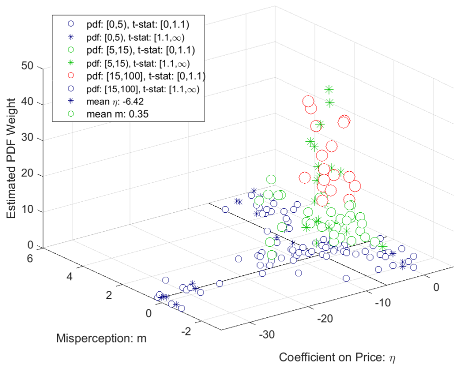


(e) Model IV τ Heterogeneous; $f(\eta, m)$

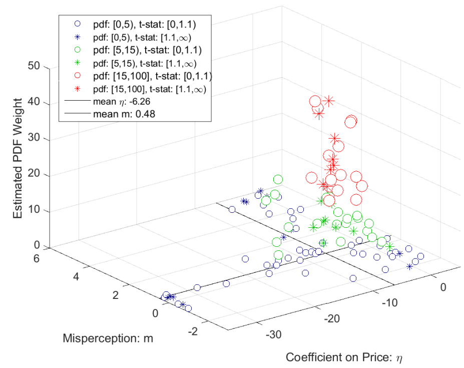


(f) Model IV τ Heterogeneous; $f(m)$

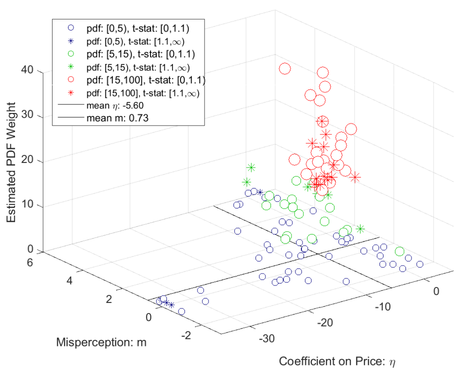
FIGURE 6. Robustness Tests: $f(\eta, m)$ and $f(m)$



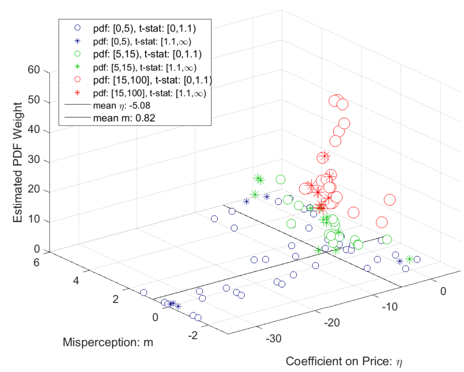
(a) Income: <$30k



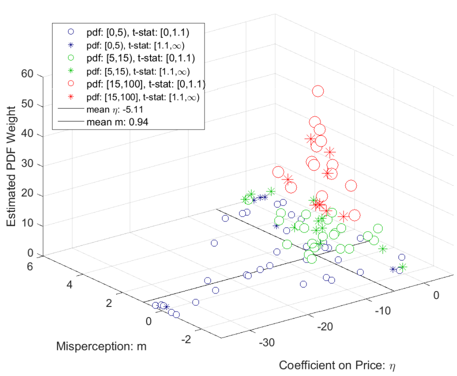
(b) Income: [$30k, $50k)



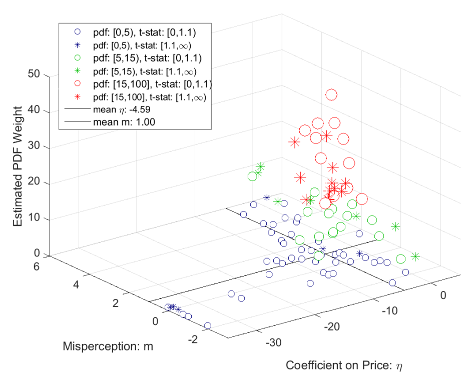
(c) Income: [$50k, $75k)



(d) Income: [$75k, $100k)

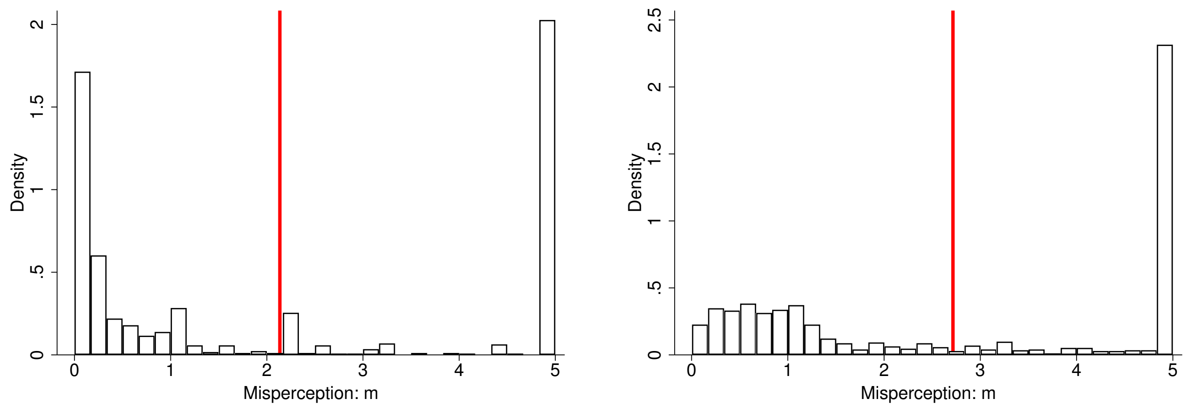


(e) Income: [$100k, $150k)



(f) Income: ≥$150k

FIGURE 7. Joint Probability Distribution for η and m by Income Group
Notes: : .



(a) Beliefs: Refrigerator Usage

(b) Beliefs: Electricity Price

FIGURE 8. Beliefs Valuation Ratios: Refrigerator Usage and Electricity Price
Notes: The red bar identifies the mean valuation ratio.

Appendix

For Online Publication

A1. Additional Summary Statistics and Results

Table A1. OLS Regressions: Robustness Checks

Purchase Price	-0.000387***	-0.000388***	-0.000406***
	(0.0000113)	(0.0000113)	(0.0000111)
Energy Cost	-0.00616***	-0.00656***	-0.00514***
	(0.000690)	(0.000705)	(0.000592)
Misperception: m	1.3617	1.4463	1.083
Product FE	Yes	Yes	Yes
Week FE	No	Yes	No
Brand × Week FE	No	No	Yes
County × Year FE	No	Yes	Yes
County × Week FE	Yes	No	No
County × ES FE	No	Yes	Yes
County × EE-related FE	No	No	Yes
Avg. Elec Price	Cty	Cty	State
# Obs	18724395	18724395	18724395

Table A2. Matched Pairs: Summary Statistics

		Matched Paired	All Models
# Models		102	6,859
MSRP ($)		1,073	1,671
kWh/y		493	575
Adjusted Volume (Cu. Ft.)		24	27
% more efficient than minimum (%)		11	10
Year entered on market		2004.8	2004.1
Δ kWh/y	mean	-69.51	-
	std	27.63	-
	10^th Pctile	-108.00	-
	10^th Pctile	-35.00	-
Δ Elec Cost $/y	mean	-8.80	-
	std	4.57	-
	10^th Pctile	-14.08	-
	10^th Pctile	-3.34	-
Δ Price $	mean	22.54	-
	std	151.64	-
	10^th Pctile	-150.00	-
	10^th Pctile	155.00	-

A2. **Monte Carlo**

A2.1. **Performance of the Mixed Logit: One-Choice Situation Case**

We investigate the performance of the mixed logit for the case where we observe only one choice for each decision-maker. We vary three dimensions: the number of alternatives (3, 6, or 10), the number of decision-makers (1,000, 10,000, or 25,000), and the parametric distribution of preferences. The data generating process is based on the following model. Decision-makers make a discrete choice among J alternatives where the alternative-specific utility of option j for consumer i is function of the price, denoted P_j, a second attribute that may represents energy operating cost, denoted E_j, overall quality, denoted γ_j, and an idiosyncratic taste parameter, denoted ϵ_{ij}:

$$U_{ij} = \eta_i P_j + \theta_i E_j + \delta_j + \epsilon_{ij} \tag{A1}$$

The preference parameters η_i and θ_i are random coefficients that follow a specific parametric distribution. Finally, we assume that ϵ_{ij} is i.i.d. and follows a extreme value distribution. This gives rise to the mixed logit.

For the first set of scenarios, we consider that the parameters η_i and θ_i follow a bivariate normal: $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = [\mu_\eta = -1, \mu_\theta = -0.7]$ and $\boldsymbol{\Sigma} = [\Sigma_{1,1} = 0.5, \Sigma_{1,2} = \Sigma_{2,1} = 0.25, \Sigma_{2,2} = 0.75]$. We generate 100 datasets for a given number of alternatives, J, and number of decision-makers, N. The price P_j, the attribute E_j, and quality γ_j are generated with a multivariate standard normal distribution. We assume that there is no correlation between the level of each of those variables across alternatives.

For each simulated dataset, we estimate a mixed logit where the preferences are specified with a bivariate normal. We recover the estimates of the mean and the full covariance matrix via simulated maximum likelihood and use 500 draws to simulate the distribution of preferences. Results are presented in Table A3. For a given J and N, we report that average of the point estimates across the 100 datasets. A number of important patterns emerge. As the number of observations increase the bias in the estimate of the mean vector $\boldsymbol{\mu}$ quickly goes to zero. This is, however, not the case for the components of the covariance matrix: $\boldsymbol{\Sigma}$. Even for a large number of observations, e.g., $N = 25,000$, the estimated covariance matrix differs from the true data generating process. Increasing the number of alternatives helps identifying the components of the covariance matrix. This is intuitive. Adding alternatives generate more substitution events that violate the independence of irrelevant alternative (IIA) assumption, which allows to pin dow the preference parameters of the covariance matrix, which is responsible for the violation of this assumption.

For the second and third sets of scenarios, we consider the effect of mis-specifying the parametric distribution of preferences. In particular, for the data generating process, we consider that preferences are generated by a mixture of two bivariate normal: $\omega_1\mathcal{N}(\boldsymbol{\mu_1}, \boldsymbol{\Sigma_1}) + (1-\omega_1)\mathcal{N}(\boldsymbol{\mu_2}, \boldsymbol{\Sigma_2})$, where $\omega_1 = 0.7$. We first set $\boldsymbol{\mu_1} = [\mu_\eta^1 = -2.0, \mu_\theta^1 = -1.5]$ and $\boldsymbol{\Sigma_1} = [\Sigma_{1,1}^1 = 0.5, \Sigma_{1,2}^1 = \Sigma_{2,1}^1 = 0.25, \Sigma_{2,2}^1 = 0.75]$ and $\boldsymbol{\mu_2} = [\mu_\eta^2 = -0.5, \mu_\theta^2 = -0.05]$ and $\boldsymbol{\Sigma_2} = [\Sigma_{1,1}^2 = 0.15, \Sigma_{1,2}^2 = \Sigma_{2,1}^2 = -0.05, \Sigma_{2,2}^2 = 0.10]$. We then consider a mixture where the two modes have a larger difference and set $\boldsymbol{\mu_1} = [\mu_\eta^1 = -6.0, \mu_\theta^1 = -4.5]$ and $\boldsymbol{\mu_2} = [\mu_\eta^2 = -0.5, \mu_\theta^2 = -0.05]$. For this latter case, all other preference parameters definying the mixture distribution are the same.

Results are presented in A4. The main take-away is that the mis-specified model does well in recovering the mean of the mixture of distribution, but does not perform well to recover the covariance matrix. This is especially the case if the two modes of the mixture distribution are much different. Adding alternatives or observations improve the estimates, but large biases remain. Note that the mixing of two normal distribution may not result in a symmetric distribution, but the mean is always a simple weighted average of the mean of each mode. By specifying a mixed logit with a normal distribution, it is therefore difficult to capture the skewness in the resulting mixture distribution, but the mean can easily be recovered.

Table A3. Monte Carlo: Performance of the Mixed Logit for the One-Choice Situation Case

	True Value	Estimates N=1,000	N=10,000	N=25,000	% Bias N=1,000	N=10,000	N=25,000
# Alternatives: 3							
η	-1.00	-1.05	-1.02	-0.99	4.84	1.75	-0.88
θ	-0.70	-0.77	-0.70	-0.69	10.03	-0.05	-0.98
$Cov_{1,1}$	0.50	1.57	0.62	0.46	214.46	23.53	-7.91
$Cov_{2,1}$	0.25	0.30	0.11	0.19	20.64	-55.33	-24.42
$Cov_{2,2}$	0.75	2.38	0.89	0.80	217.28	18.21	6.08
# Alternatives: 6							
η	-1.00	-1.05	-1.00	-0.99	5.00	0.13	-0.52
θ	-0.70	-0.72	-0.70	-0.70	3.10	-0.12	-0.55
$Cov_{1,1}$	0.50	0.92	0.48	0.41	83.87	-3.48	-17.74
$Cov_{2,1}$	0.25	0.27	0.21	0.26	6.32	-17.61	4.47
$Cov_{2,2}$	0.75	1.11	0.69	0.71	48.63	-8.46	-4.71
# Alternatives: 10							
η	-1.00	-1.03	-1.00	-0.99	2.50	-0.06	-0.82
θ	-0.70	-0.70	-0.71	-0.70	-0.52	0.84	-0.08
$Cov_{1,1}$	0.50	0.87	0.43	0.44	74.85	-14.07	-11.55
$Cov_{2,1}$	0.25	0.19	0.25	0.24	-25.71	0.37	-2.24
$Cov_{2,2}$	0.75	1.07	0.65	0.72	42.21	-13.41	-4.31

Table A4. Monte Carlo: Performance of the Mixed Logit with Mis-specified Distribution

	True Value			Estimates			% Bias		
				N=1,000	N=10,000	N=25,000	N=1,000	N=10,000	N=25,000
	$\mathcal{N}(\boldsymbol{\mu_1}, \boldsymbol{\Sigma_1})$	$\mathcal{N}(\boldsymbol{\mu_2}, \boldsymbol{\Sigma_2})$	Mixture						
Mixture I. Small difference in means									
ω_i	0.70	0.30							
η	-2.00	-0.50	-1.55	-1.64	-1.55	-1.55	6.06	0.29	-0.31
θ	-1.50	-0.05	-1.07	-1.14	-1.05	-1.06	6.58	-1.47	-0.70
$Cov_{1,1}$	0.50	0.15	0.87	2.03	1.00	0.91	139.40	18.18	7.43
$Cov_{2,1}$	0.25	-0.05	0.62	0.51	0.49	0.58	-15.53	-19.02	-3.87
$Cov_{2,2}$	0.75	0.10	1.00	2.38	1.13	0.98	138.47	12.76	-1.63
Mixture II. Large difference in means									
ω_i	0.70	0.30							
η	-6.00	-0.50	-4.35	-4.68	-4.50	-4.48	7.64	3.47	3.08
θ	-4.50	-0.05	-3.17	-3.44	-3.27	-3.27	8.73	3.26	3.23
$Cov_{1,1}$	0.50	0.15	6.75	11.34	10.04	9.99	68.06	48.81	47.95
$Cov_{2,1}$	0.25	-0.05	5.30	8.36	7.98	7.90	57.84	50.62	49.01
$Cov_{2,2}$	0.75	0.10	4.70	9.48	7.01	6.82	101.66	49.05	45.12

Notes: The column referring to the true value of the mixture distribution corresponds to the mean and covariance matrix of a unimodal distribution fitted on the true mixture distribution of preferences. We refer to it as the true unimodal distribution. The mixed logit specified with a bivariate normal is compared to this distribution. The true values of the covariance matrix of the mixture distribution are computed numerically.

A2.2. **Performance of the FKRB Estimator**

The FKRB estimator is implemented as two-step estimator. In the first step, all the parameters, δ_j, η, and θ of the discrete choice model are estimated using a conditional logit or a mixed logit with bivariate normal for the distribution of η, and θ. In the second step, the product fixed effects, γ_j estimated in the first step are treated as data and the FKRB estimator is implemented with non-parametric joint distribution specified for η, and θ.

Table A5 compare the mean of the estimated non-parametric distribution of the parameters η and θ. In the first panel, the product fixed effects are from the conditional logit. In the second panel, the product fixed effects are from the mixed logit. The two approaches yield similar results if the true distribution of preferences follows the simple bivariate normal. However, if the true data generating process follows a more complex mixture distribution, using the product fixed effects from the conditional logit can lead to large bias in the mean estimated of η and θ. Using product fixed effects from the mixed logit reduces substantially the bias. The difference between the two approaches also impact the overall distribution. Figure A1 compared the estimated non-parametric distributions under the two approaches for the case where the true data generating process the mixture with the two modes that have a large difference in means. When the product fixed effects from the mixed logit are used, the estimated distribution fits well the true distribution and capture the two modes.

A3. **Implementation FKRB Estimator**

Forthcoming...

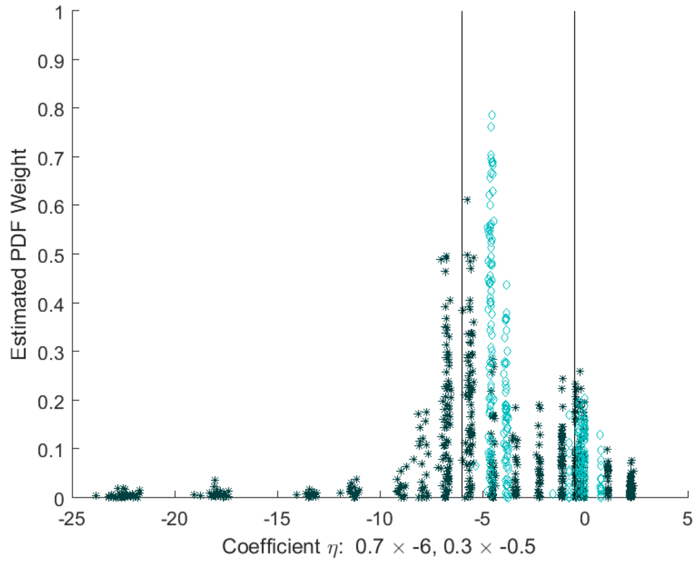
A4. **Theory: Proofs and Additional Results**

Forthcoming...

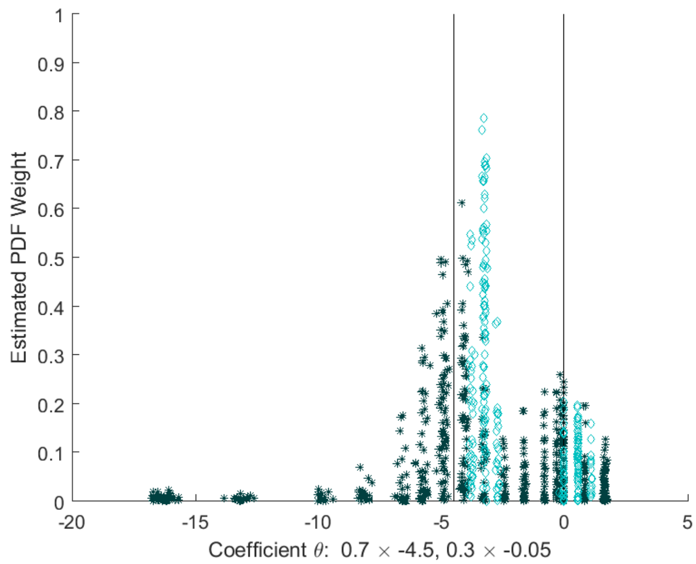
Table A5. Monte Carlo: Performance of the Mixed Logit with Mis-specified Distribution

	True Distribution	True Value	Estimates N=1,000	Estimates N=10,000	Estimates N=25,000	% Bias N=1,000	% Bias N=10,000	% Bias N=25,000
Product Fixed Effects from Conditional Logit								
η	Bivariate normal	-1.00	-1.05	-1.03	-0.99	5.17	2.78	-0.97
θ		-0.70	-0.78	-0.71	-0.70	12.12	1.13	0.27
η	Mixture I.	-1.55	-1.65	-1.54	-1.52	6.17	-0.71	-1.72
θ		-1.07	-1.13	-1.05	-1.05	6.03	-1.77	-1.54
η	Mixture II.	-4.35	-3.72	-3.65	-3.64	-14.44	-16.16	-16.40
θ	Large difference in means	-3.17	-2.71	-2.61	-2.62	-14.31	-17.40	-17.21
Product Fixed Effects from Mixed Logit								
η	Bivariate normal	-1.00	-1.12	-1.05	-1.01	12.12	5.34	1.07
θ		-0.70	-0.85	-0.73	-0.72	21.87	3.98	2.90
η	Mixture I.	-1.55	-1.76	-1.59	-1.56	13.48	2.38	0.92
θ		-1.07	-1.21	-1.08	-1.08	13.82	1.63	1.73
η	Mixture II.	-4.35	-4.99	-4.56	-4.52	14.80	4.78	3.83
θ	Large difference in means	-3.17	-3.71	-3.32	-3.29	17.07	4.83	4.05

Notes: The first panel report results where the fixed effects in the discrete choice model are first estimated using a conditional logit and then treated as data in the FKRB estimator. The second panel reports results where the fixed effects are estimated with a mixed logit where the preference parameters follow a bivariate normal. For all scenarios, the number of alternatives if fixed to 3, which means that two product fixed effects need to be estimated.



(a) Estimated pdf for η



(b) Estimated pdf for θ

FIGURE A1. Performance of FKRB Estimator for a Mixture Distribution
Notes: : Each panel shows the estimated marginal distribution of η and θ recovered from the non-parametric joint pdf. Each $\lozenge$ corresponds to the weight of the distribution estimated with the product fixed effects from the conditional logit, and each $*$ corresponds to the weight of the distribution estimated with the product fixed effects from the mixed logit. The modes of the mixture distribution of the true data generating process are depicted by the vertical lines. For all simulations, the number of alternatives is $J = 3$ and the number of observations is $N = 25,000$.