The Incidence of Coarse Certification: Evidence from the ENERGY STAR Program

Sébastien Houde*

May 4, 2025

Abstract

A coarse certification provides simple but incomplete information. Its rationale is to help consumers trade off dimensions of quality that are complex and lack salience. In imperfectly competitive markets, it may induce excess bunching at the certification requirement, crowd out quality, and facilitate price discrimination. Who will ultimately benefit from a coarse certification thus depends on the degree of market power firms can exercise as well as on consumers' sophistication in responding to such information. I use a structural econometric model of the U.S. appliance market to illustrate these insights using the ENERGY STAR certification as a case study.

JEL: D43, L13, L15, L68, Q48.

Keywords: Coarse certification, consumer attention, differentiated markets, structural estimation, energy efficiency.

---

*HEC Lausanne, e-mail: sebastien.houde@unil.ch. I would like to thank Wesley Hartmann, Tomomichi Amano, Jonathan Levin, James Sweeney, Louis-Gäetan Giraudet, Massimo Filippini, Ginger Zhe Jin, Christopher Timmins, Shanjun Li, and several seminar participants for helpful suggestions and comments. This research was part of the activities of SCCER CREST, which is financially supported by the Swiss Commission for Technology and Innovation (CTI).

1 Introduction

The role of certification is to correct informational market failures when a dimension of quality is hard to assess. In practice, certification programs established by governmental entities, nonprofit organizations, or trade associations offer a limited amount of hard information but instead provide coarse signals. For instance, a certification often takes the form of a letter-grade system, star rating, or discrete numerical scale. Coarse certifications are ubiquitous in the financial, health, real estate, food, and energy sectors, among many others.

The main rationale of a coarse certification is to provide a simple heuristic to compare products. It is often managed like a brand to increase the salience of one dimension of quality. When successful in branding, a coarse certification program then dictates the equilibrium in the market: consumers have a high willingness to pay for certified products, and firms respond by offering products that meet the certification requirement. Ultimately, it may increase the provision of the hard-to-assess dimension of quality.

However, the coarse nature of a certification can lead to unintended consequences. When a certification program complements readily available information that is accurate but complex, it may crowd out efforts to rely on more accurate information signals (Houde 2018). Taking into account the strategic behavior of firms, this in turn may induce excess bunching at the certification requirement, distort prices, and crowd out the quality that the certification aims to increase in the first place.

Who ultimately benefits from a coarse certification thus depends on the degree of market power firms can exercise as well as on the share of sophisticated consumers that respond to such an information signal. This paper illustrates these insights using the ENERGY STAR (ES) certification program as a case study. I investigate the incidence of the program with a structural econometric model of the US appliance market.

The analysis proceeds in three steps. I first develop a theoretical framework that provides intuition on why the welfare effects and the incidence of a coarse certification are a priori ambiguous. The framework formalizes that the incidence of a coarse certification depends on the degree of market power and heterogeneity in consumer sophistication. Second, I apply the framework to the energy domain to investigate the ES program, one of the most well-known environmental certifications used in the United States and Europe.[1] The goal of this program is to favor the adoption of energy-efficient products by providing simple and salient information to consumers.

[1] The ES program was first established in the United States in 1992 but since then it has been adopted in Canada, India, and several European countries.

I develop a stylized oligopoly model of the US refrigerator market and carry out a structural estimation with rich micro data. Finally, I simulate the market with and without certification to quantify the welfare effects associated with ES for different market participants and investigate different counterfactual scenarios.

In my main policy simulation, I find that ES crowds out the provision of energy efficiency and reduces welfare. Consumers are better off without certification, and this holds across income groups and regions subject to low and high electricity prices. High-income households living in regions with high electricity prices gain the most without certification, because a large share of products offered exceed the certification requirement, and these households tend to value energy efficiency more. Firms tend to benefit from the certification, but the effect on profits is small and heterogeneous. These results crucially depend on various features of the market environment. The crowding out of energy efficiency occurs because there is a large enough share of consumers that would respond to energy cost information in the absence of certification. Therefore, firms find it optimal to offer products with higher efficiency levels that largely exceed the certification requirement that was in effect. In the absence of certification, the share of products that meet the federal minimum energy efficiency standard increases, intensifying price competition in this region of the product space and benefiting consumers who value energy efficiency less. More generally, the degree of market power firms hold, the electricity prices, and the share of sophisticated consumers affect the magnitude and direction of the welfare effects. Moreover, I show that a more stringent certification requirement could have made ES welfare-improving in my context.

Although ES is one of the main federal policies used to manage energy demand and has been adopted in several countries, this paper is the first to conduct a comprehensive welfare analysis of the program accounting for firm behavior and consumer heterogeneity. Prior work on ES has focused on estimating how consumers value the certification (Eichholtz et al. 2010; Houde 2018; Walls et al. 2017; Ward et al. 2011). These studies show that consumers' willingness to pay (WTP) for ES is significant and tends to exceed the monetary value of the energy savings associated with certified products. A few studies have also investigated firms' strategic response to the certification. Allcott and Sweeney (2016) studied the behavior of sales agents selling ES-certified products and found that agents were selectively choosing to offer ES-certified products to different consumers. In Houde (2022), I documented several stylized facts pertaining to firms' product lines and pricing decisions in response to ES. I show that firms maintain a price premium for ES products that closely matches consumers' WTP. Spurlock and Fujita (2022) also shows that firms use the ES certification and mandatory energy efficiency standards to second-degree price discriminate. In the broader context of environmental certifications, Rysman et al. (2020) show that builders exploit coarse certification to vertically differentiate from competitors. Finally, Amano and Ohashi (2018) studied a Japanese minimum standard, similar in design to ES, and they found that the coarse nature of the regulation, together with the dynamic incentives set by the regulator, induces firms to retain product innovation strategically.

This paper complements a large body of work on instrument choice for energy and environmental policy. Certification programs are a popular type of information-based policy used to account for environmental externalities. Several theoretical studies have investigated issues that arise in the design of environmental certifications, such as competing labels (Fischer and Lyon 2014; Heyes and Martin 2016), consumer confusion (Harbaugh et al. 2011), and imperfect certification requirements (Mason 2011). An important takeaway from this literature is that environmental certifications are not guaranteed to improve welfare. Moreover, they may even have the unintended consequence of decreasing environmental quality (Kotchen 2006). I reach similar conclusions by focusing on the interaction between two market failures: imperfect competition and consumers' costly information acquisition, which I refer to as microfrictions. In the presence of microfrictions, the introduction of a coarse certification induces some consumers to rely on a coarse signal instead of a more accurate signal. In equilibrium, this induces firms to offer fewer products that exceed the certification requirement and can lead to an overall decrease in quality. I refer to this phenomenon as the crowding-out effect. Welfare is also affected by distortions in prices due to imperfect competition. I show that a coarse certification segments the market in one dimension of quality, which relaxes price competition and leads to higher markups. The level of coarseness is a key policy design choice. Using the model, I illustrate how introducing a second tier to the ES certification can mitigate market power. However, this does not necessarily benefit consumers or improve overall welfare. A binary scheme—optimized to induce firms to offer a high share of efficient products–may be preferable to a two-tier scheme that smooths the distribution of energy efficiency.

My work is also related to studies that have focused on the car market and investigated manufacturers' strategic responses to environmental regulations, especially fuel economy standards (e.g., Holland et al. 2009; Ito and Sallee 2018; Jacobsen 2013; Klier and Linn 2012; Knittel 2011; Reynaert and Sallee 2021; Whitefoot et al. 2017). The consensus from these studies is that mandatory minimum standards reduce profits and are dominated by market-based instruments. The present paper focuses on a different market and, more importantly, on a different use of standards. The fact that ES acts as a voluntary standard and induces innovation beyond a minimum standard is an important distinction and explains why (some) firms may benefit from such certification.

This paper contributes more broadly to the literature on certifications and information disclosure programs (Dranove and Jin 2010). An important theme in this literature is whether a certification can be informative and mitigate adverse selection problems as in Akerlof (1970). A large strand of this literature studies the behavior of sellers and buyers subject to information asymmetries, where the sellers decide to self-certify. In these models, certification acts as a signaling device that usually does not provide complete information (Stahl and Strausz 2017). As in Barahona et al. (2023), I also study a certification that is not fully informative with an equilibrium model. In my setting, the coarse information is, however, a key design decision made by the regulator. I then highlight the behavioral mechanisms by which a coarse certification—providing some relevant information—might not necessarily succeed in increasing the provision of a hard-to-assess dimension of quality, and ultimately welfare.

Finally, the paper contributes to the literature using structural behavioral econometric models to conduct empirical welfare analysis. By modeling consumers' attention allocation process to different pieces of energy information, I need to take a stand on the components of welfare that are truly experienced. My approach to quantifying welfare in this behavioral framework follows the work of Allcott (2013); Dubois et al. (2017); Keane et al. (2021) where consumers' mistakes and misperceptions are accounted for in the measurement of welfare. Two innovations that I provide to this literature is first to propose a welfare measure that accounts for latent types that differ in their degree of sophistication and, second, to account for the role of firms exploiting consumers' potential biases, a theme that has remained largely unexplored with structural behavioral models.[2]

The remainder of the paper is organized as follows. The next section presents a general framework to study the welfare effects of coarse certification in imperfectly competitive markets. Section 3 discusses the empirical setting. Sections 4 and 5 develop and estimate an oligopoly model of the US appliance market. The policy analysis is performed in Section 6, and conclusions follow in Section 7.

2 Coarse Certification and Imperfect Competition

This section presents a simplified model of imperfect competition with consumers that differ in their ability to assess one dimension of quality. The model shows two important results about the role of a coarse certification in this setting. First, a coarse certification may lower the overall provision of quality. Second, the model formalizes which regions in the product space the introduction of a coarse certification may induce products to bunch at the certification requirement or not.

[2]DellaVigna (2018) conducts an exhaustive review and shows that most of the studies using structural behavioral models do not account for the supply side.

2.1 Setup

I consider a monopolistic market for a product for which consumers have unit demand. The product is a technology with a quality dimension that is difficult for consumers to assess, because information is shrouded and hard to collect, complex and hard to process, or both. I assume that consumers are heterogeneous with respect to their ability to collect and process information. This creates a gap between decision and experienced utility. Some consumers could omit some cost information about the technology when making an adoption decision, but would later experience the actual cost of the technology. For instance, the technology could be an energy-intensive durable (e.g., a car, a refrigerator, or a television), and the hard-to-assess quality dimension could be the lifetime energy operating costs, which some consumers could ignore. However, all consumers will eventually experience those costs. In this context, the firm's strategy and equilibrium outcomes will thus be based on consumers' decision utility. Welfare, on the other hand, will be based on experienced utility.

Consumers trade off the hard-to-assess attribute, a_j, for the other dimension of quality, δ_j, and the price, p_j. The attribute a is a continuous variable representing a particular vertical quality dimension. To keep the exposition simple, I focus on the case where consumers are heterogeneous only in their ability to collect and process information about a. A share of unsophisticated consumers, denoted by $H_i(U)$, lack the sophistication to assess the attribute a and dismiss it. A share of sophisticated consumers, denoted by $H_i(I)$, value a with a marginal valuation of γ_i. The valuation of a is thus heterogeneous with a mass at zero, represented by $H_i(U)$. The decision utility of consumers of type i from purchasing product j is thus

$$U_{ij} = \begin{cases} \delta_j - p_j, & \text{with probability} \quad H_i(U) \\ \gamma_i a_j + \delta_j - p_j, & \text{with probability} \quad H_i(I) = 1 - H_i(U) \end{cases} \tag{1}$$

Underlying probabilities $H_i(U)$ and $H_i(I)$ are a process whereby various microfrictions affect how consumers evaluate attribute a.[3] In expectation, consumers' decision utility is

$$\bar{U}_{ij} = \theta_i a_j + \delta_j - p_j, \tag{2}$$

[3]Several behavioral models can provide the microfoundations to model such a process. In the empirical application, I propose a rational model of attention allocation in the spirit of Sims (2003) and Sallee (2014) to capture consumers trading off coarse information versus accurate but more costly information. In this model, heterogeneity in the cost of collecting and processing information leads to discrete types with respect to the degree of sophistication with which consumers respond to the hard-to-assess attribute. Alternative behavioral models could also be used to model this trade-off (e.g., Mullainathan et al. 2008), and other microfrictions, such as biased beliefs, could also explain why consumers would not consider or misperceive attribute a.

where $\theta_i = \gamma_i \cdot H_i(I)$. I assume that the firm has prior beliefs about the distribution of the parameter θ_i.[4] Much of the intuition can be derived for the case where the firm has prior beliefs for two segments such that $\theta_i = \{\theta_L, \theta_H\}$, where $\theta_L < \theta_H$, and π is the prior belief that consumers are of type L. The firm's prior beliefs could be determined by past market experience. For instance, this scenario could represent the case where the monopolist has learned the share of low- versus high-income consumers in the focal market. In addition, the monopolist could believe that low-income consumers tend to be more inattentive and have a low marginal valuation for the attribute a (i.e., they have a high value of $H_i(U)$, but a low value of γ_i) relative to high-income consumers. I assume that the firm strategy does not affect the parameter θ_i.[5] Finally, I assume that the cost to produce attribute a, denoted $C(a)$, is increasing and convex.

In this framework, there is an important conceptual distinction between the parameters γ and θ. Whereas γ represents preferences and captures the utility that a consumer experiences upon purchasing a product, θ represents a firm's beliefs about consumers' decision utility at the time of purchase. The difference between γ and θ thus produces a gap between decision and experienced utility—that is, the utility a consumer is expected to experience ex ante versus the utility that is actually experienced ex post (Kahneman et al. 1997).

Under this setup, there are three market failures: imperfect competition, asymmetry of information between the firm and consumers, and microfrictions, which induce different levels of consumer sophistication. The monopolist's optimal choice for the level of attribute a and the price offered to each consumer segment is the solution of the canonical screening problem of Mussa and Rosen (1978). In the present setting, the social planner's solution differs due to the gap between decision and experienced utility induced by microfrictions. In Appendix A, I provide a complete model derivation. However, most of the intuition can be obtained with a graphical approach. Panel (a) Figure 1 first shows the gap between decision and experienced utility in this model. On the y-axis, we have the level of utility, which can be measured in dollar units given that preferences are quasi-linear. On the x-axis, we have the level of quality for a_j. Consumers' experienced utility for each type starts at δ_j for $a_j = 0$ and increases at a rate depicted by the slope of the dotted line.

[4] Alternatively, instead of having beliefs about θ_i, the firm could have different beliefs about each element of θ: i.e., the parameters γ_i and $H_i(I)$. In such a case, the firm could screen consumers in multiple dimensions. The present case corresponds to a less sophisticated firm that can only screen in one dimension. The main theoretical results are thus based on heterogeneity on the parameter θ_i, and how it changes when a coarse certification is introduced. The parameter γ_i and probability $H_i()$ aim to clarify the underlying behavioral factors that would lead to variation in θ_i.

[5] The empirical model relaxes this assumption and endogenizes the latent probabilities $H_i()$, which are a function of the choice set offered by the firms.

Inattention implies a downward shift in the slopes,[6] which the monopolist will use to determine the optimal location of a_j in the price-quality dimensions. Panel (b) adds the cost of providing a_j and shows the optimal provision of a_j for both types. The social planner will address the three market failures at once by setting the level of a_i for each consumer type such that the marginal cost of producing attribute a equals the true (i.e., the experienced) marginal valuation: $C'(a_i) = \gamma_i$. The social planner then determines a price for this attribute level that maximizes the surplus of each consumer. It is thus as if the social planner assigns a product to each consumer type. Note that in Panel (b), I have assumed that the social planner will prefer giving all the surplus to the consumers and setting the profits to zero.

Panel (c) shows the monopolist's strategy. The optimal location of a_j in the price-quality dimensions is based on her beliefs about consumers' decision utility, in particular, the realization of θ_H and θ_L. The firm can distort quality to screen consumers,[7] and it sets $C'(a_H) = \theta_H$ and $C'(a_L) = \frac{\theta_L-(1-\pi)\theta_H}{\pi}$. Because $\theta_H < \gamma_H$ and the cost is increasing with a, a_H will be under-provided relative to the social optimum. This is a notable difference from the classic screening problem, where only quality for the lower type is distorted. In the present case, the monopolist will thus underprovide quality at both ends of the quality spectrum (Panel (d), Figure 1). The monopolist reduces the quality of the product offered to the low type because of the downward shift in decision utility and to relax the incentive compatibility constraint of the high type. The relative degree of inattention between the low and high types will thus determine how much distortion there is at the lower end of the quality spectrum. At the other end of the quality spectrum, the degree of inattention of the high type solely determines the distortion in quality. However, the relative degree of inattention of both types determines the reduction in price for the high-quality product.

2.2 Coarse Certification

I model a coarse certification as follows. I assume it provides a simple and salient signal about the value of a. The coarse information signal influences the purchase decision via two mechanisms. First, the certification can be informative and provide a heuristic to assess the value of a. In this case, consumers might form beliefs about the average value of a for certified and non-certified products. Consumers may thus value certified products, which I note $D_j = 1$, as a function of the conditional means of a: $E[a|D_j] = \mathring{\tau}_i$. Second, the certification, similar to some types of advertising,

[6] As shown on the figure, the downward shift for the two dotted lines is about the same for both types. This is for illustrative purposes only and without loss of generality. Inattention could be very heterogeneous across types.

[7] If θ_L is small relative to θ_H, there might not be a solution for the optimal level offered to type L consumers. In this case, the firm might simply set a at its minimum $a_L^* = \underline{a}$, or offer only one product (pooling equilibrium). Therefore, there are two possible equilibria: a separating equilibrium and a pooling equilibrium.

might be persuasive (Bagwell 2007) and influence preferences directly, irrespective of the level of a. I will be agnostic on the nature of the behavioral mechanisms leading to persuasion and assume that if a is a desirable attribute, $\tilde{\tau}_i \geq 0$ is the persuasion effect.

Consumers can now fall into three categories: with probability $\hat{H}_i(C)$, a consumer will rely on the coarse certification; with probability $\hat{H}_i(U)$, a consumer will remain unsophisticated; and with probability $\hat{H}_i(I)$, a consumer will be sophisticated. In the presence of the certification, the utility of consumer i from purchasing product j becomes

$$U_{ij} = \begin{cases} \delta_j - p_j, & \text{with probability } \hat{H}_i(U) \\ \mathring{\tau}_i D_j + \tilde{\tau}_i D_j + \delta_j - p_j, & \text{with probability } \hat{H}_i(C) \\ \gamma_i a_j + \delta_j - p_j, & \text{with probability } \hat{H}_i(I) = 1 - \hat{H}_i(U) - \hat{H}_i(C) \end{cases} \tag{3}$$

I assume that the firm has beliefs about $\tau_i = (\mathring{\tau}_i + \tilde{\tau}_i) \cdot \hat{H}_i(C)$, but not specifically about $\mathring{\tau}_i$ and $\tilde{\tau}_i$. From the standpoint of the firm, consumers' expected decision utility can now be expressed as

$$U_{ij} = \hat{\theta}_i a_j + \tau_i D_j + \delta_j - p_j \tag{4}$$

where D_j takes the value of one if product j is certified and zero otherwise, and $\hat{\theta}_i = \gamma_i \cdot (1 - \hat{H}_i(U) - \hat{H}_i(C))$.

In Equation 4, the coarse certification affects utility by creating a discrete increase of size τ_i in the willingness to pay for certified technologies and by changing the marginal valuation of the hard-to-assess attribute from θ_i to $\hat{\theta}_i$. Note that $\theta_i \geq \hat{\theta}_i$ if the certification changes the shares of unsophisticated and sophisticated consumers as follows: $H_i(U)$ becomes $\hat{H}_i(U) \leq H_i(U)$ and $H_i(I)$ becomes $\hat{H}_i(I) \leq H_i(I)$. The first inequality is intuitive. Some consumers might find it too difficult to fully assess attribute a but can process the coarse information signal. Therefore, when a certification is introduced, some consumers who used to dismiss a completely now account for it using the heuristic provided by the certification. The second inequality captures the fact that a coarse certification might allow some consumers to economize on the efforts required to assess a. In this case, the certification crowds out the share of better-informed consumers toward a share of consumers that rely on the simpler but coarser information signal (Houde 2018).[8] When this latter phenomenon occurs, the following inequality holds: $\theta_i > \hat{\theta}_i$, which has important implications for determining the equilibrium outcomes under a coarse certification.

[8] In this section, I do not fully endogenize the attention allocation process. The case where the shares of unsophisticated and sophisticated consumers decrease when a coarse certification is introduced encompasses several behavioral models, including a model of rational inattention as in Sallee (2014), which forms the basis of the demand model in the empirical exercise.

Figure 2 (Panel a) illustrates the two mechanisms. As before, the dotted line represents the experience utility. At the certification requirement, denoted by the vertical line, the marginal valuation of a increases by $\tilde{\tau}_H$, which is the experienced component of utility associated with the certification itself. Decision utility is distorted in two ways compared to a market without certification. First, there is the discontinuous jump in utility at the certification requirement. Second, the marginal valuation of a decreases: θ_i becomes $\hat{\theta}_i$, where $\theta_i > \hat{\theta}_i$, due to the crowding out of informed consumers. In a separating equilibrium, the level of a offered to the high type, which is determined by the equality $\hat{\theta}_i = C'(a_H^*)$, is thus decreasing (Panel d). When introducing the coarse certification reduces the overall share of informed consumers, the firm reduces quality at the higher end of the spectrum. Readily, we can see why the certification might have an ambiguous impact on the overall provision of quality. If the crowding-out of high quality is important enough, it could dominate quality improvement of the low-quality model.

The large discontinuity at the certification requirement also implies that the monopolist might want to set $a_H^* = a^C$, although $\hat{\theta}_i < C'(a^C)$. As I show formally in Appendix A, the impact of bunching on the overall provision of a is, however, uncertain relative to a market without certification. For instance, if without certification $a_H^* < a^C$, but with certification $a_H^* = a^C$, the coarse certification increases the provision of a. But it is possible that without certification, $a_H^* > a^C$, and the certification induces $a_H^* = a^C$. This last case is illustrated in Panel d of Figure 2—the bunching at the certification crowds out the overall provision of a. The important takeaway here is that bunching at a coarse certification can arise from either side of the certification requirement. Therefore, empirical evidence of bunching at the requirement cannot be taken as evidence that the certification is successful in improving quality.

These results also have important implications for deciding the stringency of a given certification requirement. Starting with a market not subject to a certification, it is hard to predict a priori which products will bunch or not at a specific requirement level. The only general predictions are (1) that a coarse certification will reduce the share of low and high-quality products in favor of products that just meet the requirement, and (2) products closer to the newly introduced requirement are more likely to be affected.

Finally, one subtlety in determining whether bunching at the certification is socially desirable is the interpretation of the persuasion effect: $\tilde{\tau}$. Should it be considered as a preference or as a bias? If $\tilde{\tau}$ is considered a bias that influences consumers' decisions but should not be accounted for in social welfare, bunching at a^C is not socially desirable (Panel b). On the other hand, if $\tilde{\tau}$ is considered as preferences and represents the utility gains induced by the certification that are truly experienced, the social planner might find it optimal to set quality at the certification requirement (Panel c).

2.3 Discussions

In the empirical model, we consider a setting with multiple firms that offer several products. The mechanisms highlighted above will translate in this oligopoly setting, but the impact of competition should decrease price and quality distortion (Stole 1995). In this general setting, it is harder to obtain tractable theoretical results. It is, however, useful to highlight two distinct mechanisms associated with certification and imperfect competition that would arise: the segmentation effect and differentiation effect (Bonroy and Constantatos 2014).

The segmentation effect arises when the certification induces a separating equilibrium where both certified and non-certified products are offered, which creates localized markets in the product space. As the density of the products decreases in the different dimensions of the characteristic space as a result of the segmentation, this softens price competition and enables firms to exercise more market power. A coarse certification should affect prices via the segmentation effect as the certification requirement creates a focal point to segment the market in the dimension of the hard-to-assess attribute.

The differentiation effect arises because of the heterogeneous impact of the certification on consumers' valuations of a. Formally, the asymmetric impact of the certification across consumer types will relax or tighten the set of incentive compatibility constraints faced by firms, which will ultimately impact the equilibrium prices. For instance, the fact that some consumer types might respond strongly to a coarse certification–that is, have a large latent share $H_i(C)$ together with a large τ_i–might facilitate price differentiation and lead to higher markups.

Whereas the segmentation and differentiation effects are closely interrelated, the former is a function of the number of products in various regions of the characteristic space, and the latter is a function of the heterogeneity in how much consumers value the hard-to-assess attribute. Both of these effects, together with the extent of the crowding effect, will ultimately determine the incidence of a coarse certification among the different market participants.

To sum up, the incidence of a coarse certification in an imperfectly competitive market will hinge on three features. First, the reallocation of the hard-to-assess attribute, a, in the characteristic space. In the empirical model, it is thus crucial that the empirical model endogenizes both prices and energy efficiency—the hard-to-assess dimension of quality in this setting. Second, the degree of heterogeneity in consumers' sophistication and how this changes due to a certification. This is the second important feature of the empirical model—I use a demand model where consumers allocate their attention to different pieces of energy information. Third, the market structure and, in particular, how firms determine their markups. The empirical model accounts for this by estimating product-specific cost functions, which allows me to endogenize markups.

Overall, the welfare effects of a coarse certification would depend on the distortions in prices and quality. There is also another element that could impact welfare, which is the information acquisition cost. Whether these costs are truly experienced or only impact decisions is debatable. The economic literature on salience (e.g., Chetty et al. 2009; Farhi and Gabaix 2020) tends to exclude these costs from experienced utility. Firms' strategies and the distortions I have highlighted above are, however, not affected by how we define these costs.

3 Empirical Setting: The US Refrigerator Market

I focus on the US refrigerator market, which offers several advantages for studying the ES program. First, this market is subject to an array of government policies that interact with ES, which provides both credible variation for the estimation and a relevant institutional context. Second, refrigerators are one of the few energy-intensive durables that have large energy operating costs, but for which the utilization decision may not need to be explicitly modeled. The fact that the (unobserved) utilization is likely to be idiosyncratic and not systematically correlated with the purchase decision facilitates the identification of preferences related to energy efficiency. Third, refrigerators are relatively simple technologies, which have not been subject to important innovation trends during my sample period. Again, this simplifies the estimation and identification of preferences, and notably motivates my static framework. Finally, the US refrigerator market, like for several other types of energy-intensive durables, is an oligopolistic market where the effects of imperfect competition are important (Ashenfelter et al. 2013).

3.1 Government Policies

In the United States, like in many other countries, government agencies have established certification programs to favor the adoption of energy-efficient appliances. The main rationale of such programs is that energy efficiency, in particular, the lifetime energy operating cost of an appliance, is difficult to assess and not fully salient to consumers. A certification that provides a simple and salient information signal can then play a role in helping consumers easily identify the most energy-efficient products on the marketplace, and ultimately induce firms to offer and advertise such products.

The ENERGY STAR (ES) program—a voluntary certification that was first established by the US Environmental Protection Agency (EPA) in 1992—exemplifies how these programs work. The EPA sets certification requirements, and products that meet or exceed the requirements can be certified with the ES label (Figure 6(a)). The label consists of a simple logo that does not contain technical information. The certification requirements for ES are usually binary—products are ES labelled or not. Therefore, the ES label only provides a coarse signal about energy efficiency. In Japan, China, India, and Europe, the design of energy labels relies on a similar approach, where a coarse star or letter grade system is used to provide information about energy operating costs.

In the US refrigerator market, technical information is also provided to consumers by the EnergyGuide label (Figure 6(d)). Unlike the ES program, EnergyGuide is a mandatory labeling program that provides detailed model-specific information about energy operating costs. In this context, consumers thus face two pieces of information to account for energy operating costs. In Houde (2018), I have shown that, although the two energy labels were designed to complement each other, they are in fact substitutes. In particular, consumers who tend to rely on the ES certification do not rely on EnergyGuide and vice versa. A significant share of consumers also appears not to rely on either of these pieces of information. I have also shown that consumers who rely on ES tend to value the ES label beyond the average energy savings determined by the certification requirement. The high willingness to pay for the certification suggests that the ES label may affect preferences directly by providing warm glow and conformity with social norms, or by enacting purely altruistic motives. It is also possible that the label biases the perception of quality. For instance, consumers might believe, wrongfully, that certified products are of higher quality, a phenomenon referred to as the halo effect (Boatwright et al. 2008). Altogether, these findings point toward the existence of different types of consumers in this market that differ in the degree of sophistication in accounting for energy information.

Apart from informing consumers, the ES certification program also plays an important role in the design of energy efficiency subsidies. In the United States, there exist several consumer rebate programs that explicitly target the adoption of ES-certified appliances. The effect of these incentives on consumers' purchase behavior tends to be highly heterogeneous across different segments of the population and program designs (Houde and Aldy 2017).

3.2 Market Structure

The US refrigerator market has an oligopolistic market structure dominated by three manufacturers: Electrolux, General Electric (GE), and Whirpool. Several mergers and acquisitions that have taken place since the early eighties, culminating with Whirlpool's acquisition of Maytag in 2006, have led to a concentrated market. In 2008, the three dominant manufacturers held about 85% of the market share for full-size refrigerators (Table B.1, Appendix B). Since then, a number of events may have led to an increase in competition. In 2016, the GE appliance business was acquired by the Chinese manufacturer Haier, while Korean appliance manufacturers have steadily gained market shares.

A particular institutional feature of the US appliance market is that manufacturers compete under various brand names, and some dominant brands, such as Kenmore, are not owned by a particular manufacturer but are sold exclusively by a national retailer. This feature of the market is believed to be important in determining the degree of market power manufacturers ultimately hold. The distribution of products across brands is still, however, fairly concentrated, especially after the Whirlpool-Maytag merger; most products were offered by the major brands associated with the top three manufacturers and Kenmore (Table B.1, Appendix B).

3.3 Data

The empirical investigation focuses on the US market for full-size refrigerators from 2008 to 2011. The primary data source consists of point-of-sale data provided by a national appliance retailer. The data are disaggregated at the transaction level and comprise the whole universe of transactions where one or more refrigerators were bought. Each row in the dataset consists of a transaction and contains information on the price and taxes paid by consumers, the manufacturer price paid by the retailer, the location of the store where each purchase was made, and the manufacturer model number. I used the manufacturer model number to match the transaction data with detailed attribute information, which includes manufacturers' reported yearly energy usage, ES certification, size, color, door design, brand, and manufacturer, in addition to several other attributes. The attribute data contain information about all refrigerator models the retailer offered during 1998-2011. Only a subset of these models was observed during 2008-2011 and thus used for the demand estimation. For approximately 40% of the transactions, I also observe consumer-specific demographic information, which includes income level, education, single vs. multifamily housing, owner vs. renter status, family size, age of the head of the household, and political orientation. The retailer collected demographic information using a data aggregator's services and matched it with each transaction whenever possible.

To complement the data from the national retailer, I also collected information about local electricity prices from the Energy Information Administration (Form EIA-861) and rebates (DSIRE database). For the demand estimation, I constructed local averages at the county level for both variables and imputed the county of residence for each consumer, assuming that a consumer lived in the same county as the store where the purchase was made.

Finally, I used data from the Federal Trade Commission and the EPA to determine which refrigerator models were on the market during the period 2008-2011. I use this information in the supply-side estimation to construct a representative choice set of the US refrigerator market during that period.

4 Econometric Model

I characterize the appliance market with a static multi-product oligopoly model where firms strategically determine the energy efficiency level and the price of each product they offer. The model aims to represent a medium-run equilibrium in prices and product lines. The model makes a number of simplifications to ensure tractability. Namely, the decisions to enter and exit the market, and to determine the size of product lines and the quality of non-energy attributes, are taken as given. Moreover, the vertical structure of the supply-side abstracts from the strategic interactions between manufacturers, brands, and retailers. Although stylized, the model has the key ingredients to illustrate and quantify the mechanisms outlined in Section 2, which allows me to determine the incidence of the ES program. Firms are strategic, and prices and energy efficiency are endogenized. The costs of producing more energy-efficient refrigerators are estimated, which allows estimating markups and the impact of ES on profits. Finally, the demand side is modeled as in Houde (2018), where consumers are heterogeneous along two dimensions: income, which is observable, and sophistication in their ability to process energy information. This rich demand model allows for simulating scenarios with and without certification and determining how consumers will reallocate in different types.

4.1 Supply

I model the vertical structure of the industry in a stylized way by abstracting the strategic interactions between manufacturers, national retailers, and local store managers. Each firm represents a brand that can decide to reallocate a pre-determined set of products in the energy efficiency space and the price of each product. I model the behavior of profit-maximizing brand managers instead of manufacturers because, in this market, several manufacturers offer similar products under different brand names. However, I omit manufacturers' strategic decisions to rebrand their products in this stylized model.[9] Moreover, I fix quality in the non-energy dimension and the choice set, i.e., the number of products each brand offers across retail stores. Relaxing those assumptions and, more generally, having a richer representation of the vertical structure should enable more opportunities for firms to exercise market power. The model is thus conservative concerning the role of imperfect competition and firms' ability to extract high markups by exploiting the certification. Empirically, it will nonetheless play an important role.

Brand managers' unit costs are the manufacturing prices they pay manufacturers and the retail costs. I assume that the manufacturing prices equal the manufacturers' marginal cost of producing one unit of a particular refrigerator model. This particular assumption encompasses two different vertical market structures. First, this corresponds to a scenario where the retailer, brand manager, and manufacturer, respectively selling, marketing, and manufacturing a model, are fully integrated (Berto Villas-Boas 2007). Alternatively, it also corresponds to a scenario where a two-part fee structure is present in the vertical contract relationship with the retailer, and the retailer has the bargaining power (Bonnet and Dubois 2010).[10]

A second important feature of the appliance market is that within a relatively short time, firms can change the energy efficiency level of their products with little effect on their overall design. It systematically occurred during the previous revisions in the ES requirement; manufacturers systematically managed to offer more energy-efficient products that were otherwise similar to previous generations. I take this as evidence that the cost of providing energy efficiency is separable from the cost of providing other attributes. I will further assume that the manufacturing prices reflect this assumption. It is a strong assumption, but it allows me to make a more precise estimation of the cost function. As I show below, the data are not rich enough to estimate a more flexible cost function. As a robustness check, I will relax those assumptions in a set of simulations.

Under these assumptions, consider that there are K brands, and brand manager k offers J_k appliance models. Brand manager k maximizes profits by choosing the energy-efficient levels, the vector $f_k = \{f_{k1}, ..., f_{kJ_k}\}$, and the prices, the vector $p_k = \{p_{k1}, ..., p_{kJ_k}\}$, of his J_k models, taking the actions of rival firms as given. Firms face a population of heterogeneous consumers in which the demand for each product is $Q_{kj}(f, p)$, and depends on all energy efficiency levels and prices

[9] This simplification mainly impacts Kenmore, a major brand that offers appliances produced by various manufacturers. The two Korean manufacturers, LG and Samsung are considered as a single brand. The other brands, except for the generic brand, are each associated with a single manufacturer.

[10] It is also possible that the observed manufacturing prices include manufacturers' markups. Another interpretation of the proposed approach is that I keep these markups constant across the counterfactuals. Again, this assumption means that I am conservative concerning the role of market power in the simulations.

($f = \{f_1, ..., f_K\}$ and $p = \{p_1, ..., p_K\}$). The problem of brand manager k consists of solving:

$$\max_{\substack{f_k=\{f_{k1},...,f_{kJ_k}\},\\ p_k=\{p_{k1},...,p_{kJ_k}\}}} = \sum_{j=1}^{J_k} (p_{kj} - c^w_{kj}(f_{kj}) - c^r_j) \cdot Q_{kj}(f, p)$$

where $c^w_{kj}(f_{kj})$ is the manufacturing cost of model j offered by brand k that varies as a function of the energy efficiency level. The term c^r_j represents a model-specific unit retail cost, which may capture advertising expenses, inventory costs, or warranty liabilities but do not vary with the energy efficiency level offered.

4.2 Demand

The demand model follows the setup in Section 2 and provides a framework to represent heterogeneity in how consumers pay attention to energy efficiency—the hard-to-assess attribute in this context. At the heart of the model is heterogeneity in the costs of collecting and processing energy information, which leads to different consumer types that differ with respect to how energy efficiency influences their purchase decisions. In particular, these costs rationalize why some consumers either dismiss the energy efficiency attribute or rely on ES; although accurate, more complex information about energy costs is readily available in this decision environment. The empirical model rationalizes observed decisions; it is thus a model of decision utility. In the policy section, I show how this model can also be used to quantify welfare based on experienced utility.

I model the purchase decision as a two-step process, where consumers first select the amount of energy information they want to collect and process, and then make a purchase decision. The decision to collect and process energy information is not observed and is thus treated as a latent decision. The model is a function of variables that vary at the household level and across regions (zip codes) and time (weeks), which I denote with the subscripts i, r, and t, respectively. The choice model takes the form of a discrete latent class model:

$$Q_{irtj} = \sum_{e=\{U,ES,I\}} H_{irt}(e) \cdot M_{irtj}(e), \tag{5}$$

where e represents the level of knowledge about energy costs each consumer acquires. Consumers fall into three mutually exclusive categories. They can be uninformed ($e = U$). In such a case, they will not know the energy cost of each product and the meaning of the ES certification. They can be knowledgeable about ES ($e = ES$) but not about the exact energy cost of each product. Finally, they can be fully informed ($e = I$) and know the energy cost of each product in their choice set. The term $H_{irt}(e)$ is the latent probability that consumer i acquires knowledge e, and $M_{irtj}(e)$ is the choice probability conditional on the level of knowledge e. The latent probability for each type represents the prevalence of each type in the population. In particular, the variables that enter those probabilities represent firms' beliefs about the factors that influence the likelihood that a consumer becomes sophisticated or not. Firms could have formed these beliefs through market experience, for instance. The term Q_{irtj} is the individual choice probability for product j, which I bring to the data.

I estimate the model with large subsamples of transaction-level data. The retailer assigns each transaction a unique identifier used to match demographic data. I compute the choice probabilities at household level i, corresponding to a combination of observable demographic characteristics. The subscripts r and t in the choice probabilities highlight the variation across regions (zip codes) and time (weeks).[11] The alternative-specific utilities that represent the decision utility for each type e enter the conditional choice probabilities denoted $M_{irtj}(e)$. For the empirical model, the alternative-specific decision utility of each type is

$$\begin{aligned}
&\text{e=I:} \quad U^{I}_{ijrt} = -\eta P_{jrt} + \delta_j + \psi R_{rt} \cdot D_{jt} - \theta C_{jr} + \epsilon^{I}_{ijrt} \\
&\text{e=ES:} \quad U^{ES}_{ijrt} = -\eta P_{jrt} + \delta_j + \psi R_{rt} \cdot D_{jt} + \tau D_{jt} + \theta ESavings_r \cdot D_{jt} + \epsilon^{ES}_{ijrt} \\
&\text{e=U:} \quad U^{U}_{ijrt} = -\eta P_{jrt} + \delta_j + \epsilon^{U}_{ijrt},
\end{aligned} \tag{6}$$

where P_{jrt} is the price, δ_j is the quality of the product, R_{rt} is the rebate amount offered for ES products, and D_{jt} takes the value one if product j is certified ES at time t and zero otherwise. The difference in alternative-specific utility for informed consumers ($e = I$) and consumers relying on ES ($e = ES$) is twofold. Informed consumers consider an accurate measure of annual energy operating costs, the variable C_{jr}, which is the product of the local electricity price, the county average in region r, and the manufacturer's reported annual electricity usage for model j. If consumers rely on ES ($e = ES$), they instead compute the average energy cost savings associated with the certification, the variable $ESavings_r$, which is the difference between the average annual electricity usage of certified products and non-certified products multiplied by the local electricity price. The second difference is that for $e = ES$, the ES label itself could influence the decision, where the parameter τ captures the behavioral response to the label. The labeling effect could capture preferences for green goods, warm glow, and various behavioral biases induced by the certification. For uninformed consumers (e=U), I assume that they dismiss all information related to energy use; they dismiss

[11] The choice probabilities for two households with the same demographic variables vary across local markets (i.e., zip code-week).

energy operating costs and are also unaware of the rebates offered for ES-certified products. Finally, ϵ_{ijrt} is an idiosyncratic taste parameter. Assuming that the idiosyncratic taste parameters ϵ are extreme value distributed, the probabilities $M_{irtj}(e)$ take the form of a multinomial logit.

In Houde (2018), I discuss the microfoundations to model the latent probabilities: $H_{irt}(e)$. In a nutshell, the costs of collecting and processing information, together with a fixed number of decision strategies to account for energy efficiency in the purchase decision, give rise to different consumer types. The latent probabilities represent the likelihood of being one of the three types. I characterize them with a parametric form. I use a multinomial logit, which can be motivated by idiosyncratic attention costs that follow an i.i.d. extreme value distribution. I also introduce variables in the latent probabilities that influence consumers' decisions to collect energy information in a model of rational attention allocation (Sallee 2014). In addition to a vector of household demographics, the probabilities for the present estimation include variables that characterize the choice set in each local market (zip code r and in a week t). Specifically, the mean and variance in electricity costs for all products offered, the mean and variance of the share of ES products offered, and the number of products in the choice set in a given region. The latent probabilities $H_{irt}(e)$ are important for simulating the counterfactual scenario without ES. When the certification is not in effect, these probabilities determine how consumers are "reallocated" between the uninformed and the informed type. The demographics and the variables that summarize the choice set generate variation in the likelihood that ES consumers sort into one of the two other types. Note that in the presence of the certification, the total share of perfectly uninformed *and* informed consumers will always be smaller relative to a market without certification. When I remove the certification, some consumers will become uninformed, and others will become informed. This feature of the demand model is true for any model of rational attention allocation Houde (2018) and is essential to capture the possible crowding out of quality predicted by the theoretical model.

4.3 Equilibrium

The empirical model follows the theoretical model presented in Section 2. Brand managers strategically determine prices and the provision of energy efficiency. However, they do not know the type of each consumer and only know aggregate demand, Q_{kj}, which corresponds to their beliefs about demand based on consumers' decision utility as in Section 2. Brand managers will thus attempt to screen consumers by distorting prices and energy efficiency.

Formally, the Nash equilibrium of the game is given by the vectors f^* and p^* that solve a system of $4 \times (J_1 + J_2 + ...J_K)$ equations. The discontinuity at the certification requirement implies that it may not be optimal to equate the marginal cost of providing energy efficiency with the marginal valuation. In the presence of ES, firms' strategies then become a discrete-continuous choice where firms must decide whether or not to bunch at the certification requirement and which price to set. The existence and uniqueness of equilibrium in this game are not guaranteed.

5 Estimation

The focus of this section is on cost estimation, especially the identification of the marginal cost of providing energy efficiency. I developed and estimated the demand model in Houde (2018). Below, I provide a succinct overview of the demand estimation and refer the readers to Appendix C for further details.

5.1 Marginal Cost of Providing Energy Efficiency

The first goal of the cost estimation is to identify the marginal cost of providing energy efficiency, which allows me to endogenize product lines in this dimension. One challenge in identifying this cost is that the first-order conditions of the oligopolistic game with respect to energy efficiency levels are not well-defined due to the presence of the coarse ES certification. My identification strategy takes advantage of an institutional feature of the US refrigerator market that allows me to recover the marginal cost with minimal assumptions about the nature of the strategic interaction between firms.

Refrigerator manufacturers commonly offer product lines that consist of a group of three to ten refrigerator models with a similar design, such as the size and door style (top freezer, side-by-side, bottom-freezer), but that differ with respect to less prominent attributes, such as the ice-maker option, the finish option (stainless or not), the color, and, in some cases, the energy efficiency levels. In some instances, it is possible to observe different refrigerator models, within a given product line, that are identical along all dimensions of quality, except their energy efficiency levels. When it occurs, one model typically meets that ES certification requirement, or a previous requirement, and another just meets the minimum standards. These product line decisions are consistent with a screening equilibrium where firms use energy efficiency to differentiate their products, and are also induced by the way the ES certification requirement is revised. More stringent ES requirements do not follow a pre-determined schedule and are usually announced only one year in advance. Manufacturers must then adapt quickly to a change in the ES requirement. In practice, they often do so by making small incremental changes to their product design to achieve energy efficiency improvements.

For instance, in 2004 and 2008, the ES requirement for full-size refrigerators was adjusted to become more stringent—prior to 2004, it was set at 10% more efficient than the minimum standard; for the 2004-2008 period, it was 15%; and it became 20% after April of 2008. Manufacturers adjusted their product lines quickly in response to these revisions. In particular, we observe that following the revision of the ES requirement in 2004 and 2008, manufacturers responded not only by offering new models that met the revised standard but by discontinuing models that were decertified (Houde 2022). The entry and exit of models around revision periods led to several instances of product lines where manufacturers offered the exact same models but differentiated only in the energy dimension. In my sample, I was able to identify 51 identical pairs of refrigerator models that differ only with respect to their annual electricity consumption—a measure of energy efficiency reported by manufacturers. To identify those identical pairs, I first used detailed attribute data to find product lines offered by the same brand, where models were of the same size,[12] door style, door material (stainless or not), ice-maker option, defrost technology, air filtration system, color, and door handle type. For each of those pairs, information from online marketplaces was also collected to compare whether all listed attributes, except for energy use, were identical. After this process, the remaining sample contains 102 refrigerator models that could be paired with an identical model. Note that within each pair, the year that a specific model entered the market may differ.

Table 1 provides summary statistics on these paired refrigerators and compares them to the overall sample of models I observe in the retailer's data. On average, these refrigerator models tend to be cheaper, smaller, and more energy-efficient than the average refrigerator model offered on the market.

For all paired refrigerator models ($N = 102$), I simply exploit variation in energy efficiency level within pair group ($G = 51$) together with the fact that I observe the manufacturer price to identify the marginal cost of providing energy efficiency. I assume that manufacturers' prices correspond to manufacturers' marginal unit costs and estimate the marginal cost by regressing the log of a manufacturer's price on a pair fixed effect, year-of-market-entry dummies interacted with brand dummies, and a proxy for energy efficiency:

$$ln(price_{j,r}^{manuf}) = \alpha + \gamma_{j,j'} + Y_j \times Brand_j + \phi Efficiency_j + \epsilon_{j,r}, \tag{7}$$

where $\gamma_{j,j'}$ is a pair fixed effect that is common to the paired refrigerator models j and j', and Y_j and $Brand_j$ are dummy variables for the year refrigerator j entered the market and its brand, respectively. These year-brand fixed effects account for various temporal shocks that might have

[12] The size attribute that I used includes a measure of freezer and refrigerator size. The height, width, and depth were also taken into account.

affected the manufacturing process and thus prices at the moment a model entered the market. For the proxy for energy efficiency, I use a functional form where energy efficiency is defined as the inverse of the annual electricity consumption. I thus expect a positive coefficient for the estimate of the parameter ϕ, which will capture that more efficient models are costlier to produce.

For my preferred estimator, the parameter ϕ has a value of 191.1 (Table 1). This estimate implies that the manufacturer price of a refrigerator model consuming 550 kWh/year will increase by 9.1% to meet the ES requirement, which corresponds to a cost elasticity with respect to energy efficiency of 0.45. This estimate is robust to different specifications: the controls for year-of-market-entry and brand have small effects.

5.2 Retail Costs

In addition to the manufacturing costs, brand managers are also facing various costs associated with retailing large appliances. These latter costs may include advertising, transportation and inventory, and warranty. Part of these retail costs might be sunk and fixed but they may also vary with the quantity sold. For instance, online advertising expenses in this market vary with demand due to the fact that retailers effectively pay for AdWords and clicks, which are correlated with the quantity sold. I estimate the retail unit costs using the assumption that firms are profit maximizing and prices are set strategically. As it is standard in the literature, I use the first-order conditions of the pricing problem to recover the cost estimates. In particular, I use the following system of equations:

$$Q_{kl}(f^*, p^*) + \sum_{j=1}^{J_k} (p^*_{kj} - c^w_{kj}(f^*_{kj}) - c^r_j) \cdot \frac{\partial Q_{kj}(f^*, p^*)}{\partial p^*_{kl}} = 0, \forall f_{kl} \tag{8}$$

where both demand, $Q(\cdot)$, and the manufacturer costs, c^w, are taken as given, and I solve for c^r. In this market, it is not realistic to assume that the unit retail costs vary systematically across refrigerator models. For instance, we should expect that a model offered by the same brand and of the same size, should incur similar advertising expenses, have similar transportation and inventory costs, and face the same warranty liabilities. Therefore, I restrict unit retail costs to vary only along key dimensions of refrigerators that should be correlated with cost heterogeneity. In particular, I assume that they vary as a function of the brand, overall size, which is also a proxy for weight, door design, and a coarse categorical variable for average price.

To construct the empirical moments given by Equation 8, I need to characterize the market in terms of brands and products. I assume that six different brands are operating in the US refrigerator market: the brands associated with the three main manufacturers, in addition to Kenmore, the Korean brands (i.e., LG and Samsung are assumed to be one brand), and a generic brand that includes all other brands. In reality, brands compete by placing different products across appliance stores, but I do not model such strategies in the present model. To reduce the problem's dimensionality, I model the game for only one representative appliance store, which aims to represent the US refrigerator market for the year 2011.[13] I set the size of the choice set to 68, and the number of refrigerators offered by each brand manager is held constant.[14] To ensure that the choice set is representative of the US market, the distribution of the 68 products in terms of brand, style, size, and energy efficiency was selected to fit the distribution observed nationally in the year 2011.[15]

The estimation results suggest that the average unit retail cost is \$259, which leads to an average markup of 31.5% of the retail prices. There is, however, substantial variation across products. For instance, the retail costs for the biggest refrigerator models are \$77 higher relative to the smallest models. Across brands, the variation in retail costs can be as considerable as \$249, holding all attributes constant. Overall, the estimates appear realistic and are consistent with other sources. For instance, my estimated markups are slightly more conservative than those used by the Department of Energy (DOE) to conduct their 2010 national impact analysis of minimum energy efficiency standards for refrigerators. The DOE then assumed that the retail markup was 37% of the final price.

5.3 Demand

I estimate the demand model using the individual choice probabilities, Q_{ijrt}, and maximum likelihood. I allow heterogeneity in all parameters that enter the demand model with respect to income by estimating the model separately for three different income groups. Pre-estimation, I created three large subsamples of transactions randomly drawn from the universe of transactions made by households that belong to a particular income group. I distinguish between households with income

[13] I chose the year 2011 because it represents a year where firms seemed to have fully adjusted to the change in the certification requirement that occurred in 2008.

[14] The size of the choice set corresponds approximately to half the number of models offered in a store in my sample. In my sample, the average number of refrigerator models offered by a store is 129 (Table B.2, Appendix B). I set the size of the choice set to 68 for computational reasons. The supply-side estimation and policy simulation results are qualitatively similar for larger choice sets.

[15] Although the main data source used for the estimation comes from a single retailer, I was able to construct a choice set that is representative of the whole US market by using data from the Federal Trade Commission (FTC), which collects annual data for all refrigerator models offered on the market. Table C.2 (Appendix B) shows how the constructed choice set matches the FTC's model shares for 2011.

of less than \$50,000, households with income between \$50,000 and \$100,000, and households with income of more than \$100,000.

Identifying the demand model requires rich sources of variation in retail prices, energy costs, information about ES, and choice set, which are all present in the data. Next, I discuss in more detail how the different sources of variation identify each key demand parameter.

The coefficient on prices exploits the retailer's national pricing policy, particularly the pricing algorithm used to set retail prices. In Houde and Myers (2021), we show that retail prices are subject to large and frequent changes that are highly idiosyncratic, product-specific, and not correlated with demand shocks. From week to week, the retail price of a given refrigerator model changes by approximately 15%, on average. This weekly variation robustly identifies the coefficient on price once we control for product fixed effects. In this context, using credible instruments for prices has little impact on the coefficient.[16]

The identification of the coefficient on energy costs is also discussed extensively in Houde and Myers (2021), where we show that local energy electricity prices provide rich variation that allows us to distinguish preferences for energy efficiency from other dimensions of a refrigerator quality, which does not vary across time and regions. Importantly, I observe the same refrigerator models being sold at stores in different electric utility territories and across time. This allows me to control for product fixed effects, which also vary with income, and to use cross-sectional and temporal variation in county-specific average electricity prices.

The behavioral response to energy costs can also be distinguished from preferences for the ES label and rebates associated with energy efficiency because both dimensions also vary. Following the revision of the ES standard for refrigerators in April 2008, many refrigerator models lost their ES certification. Using data that cover a period before and after the revision in the standard, it is possible to observe the same refrigerator model sold at the same store, with and without the ES label. This variation in labeling identifies how the label influences consumers. For rebates, only some regions offer them, varying in magnitude across times and regions. During the sample period, there was also an extensive and temporary program to subsidize ES refrigerators, which I use to identify the response to such rebates (Houde and Aldy 2017).

Finally, the change in relative prices, product entry and exit, and the ES decertification induce substitution patterns that identify heterogeneity in the way consumers process energy information and, ultimately, the parameters that enter the latent class probabilities ($H_{irt}(e)$). Choice set

[16] In Houde and Myers (2021), we instrument for retail prices using an instrumental variable strategy similar to the one proposed by Hausman (1996). This has little impact on the results.

variation across stores is also important, facilitating the identification of the latent classes. In Houde (2018), I discuss how identification works in such a model. The gist of the argument is twofold. First, each type perceives products differently in the energy efficiency dimension, which induces specific substitution patterns. The latent probabilities aim to capture the incidence of these patterns. To illustrate, if variation in energy costs, the ES label, and rebates do not influence decisions, once we control for prices and vertical quality, all consumers will be classified as inattentive. At the other extreme, if variations in energy costs systematically induce a response that is one-to-one with the response to purchase prices, all consumers will be classified as perfectly informed. To identify the share of consumers that rely on the ES certification, note that the coarseness of the label implies congestion in the product space on each side of the certification threshold. Removing the label and product entry and exit changes the degree of congestion and thus helps the identification of this type. Second, controls for vertical quality are also important. In particular, I must make an exclusion restriction between the variables that enter the latent probabilities and the variables that define each alternative's utility level. Specifically, I assume that product fixed effects, which vary with income, capture most of the variation in vertical quality for a given product, and the other demographic variables, namely education, family size, age, and political orientation, primarily impact the attention costs.

Table C.1 (Appendix C) presents the demand estimates. I estimated the model on three separate samples, one for each income group, which are clearly labeled at the top of the table columns. Panel A presents all the coefficients that enter the conditional choice probabilities specified in Equation 6. Focusing on the price coefficients, we observe an inverse correlation between consumers' sensitivity to prices and income levels, i.e., the marginal utility of income decreases with income. Meanwhile, lower-income consumers are also less sensitive to electricity costs. This latter conclusion hinges on two different effects.

First, the coefficient on electricity costs (θ), which captures the behavioral response to electricity costs for the share of informed consumers, is smaller, in relative terms, for lower-income households. To interpret the magnitude of the estimate of the sensitivity to electricity costs across income groups, I compare η and θ and compute an implied discount rate that rationalizes how much consumers discount future electricity costs. Assuming a refrigerator lifetime of 18 years, the implied discount rate is 3% for households with an income larger than $100K, 8% for households with income between $50K and $100K, and 8% for households with an income of less than $50K.

Second, lower-income households are more likely to dismiss energy information altogether. Panel B of Table C.1 presents the estimated coefficients for the variables that enter these probabilities and, at the bottom of Panel B, the average share of consumers of each type. We observe that a significant share of consumers across all income groups are highly likely to be uninformed ($e = U$). Moreover, the probability is much higher for low-income households (45%) than high-income households (27%).

The share of consumers that pay attention to ES ($e = ES$) varies from 20% to 10% across income groups. For these consumers, the effect of the ES label is positive, relatively large, and varies across income levels. The estimate of the label effect τ^{ES} translates into a willingness to pay (τ^{ES}/η) for the certification itself that ranges from \$164 to \$430. The large willingness-to-pay estimates raise the question of whether the preferences for the ES label truly reflect consumers' preferences for certified models or biases in how consumers perceive the overall quality of certified models. As I discuss below, this is an important distinction from a welfare standpoint.

6 Policy Analysis

The main goal of the policy analysis is to create counterfactual scenarios with and without the ES certification. By comparing the market equilibrium for each of these two scenarios, I can assess the welfare changes and incidence of the program on different types of consumers, firms, and externalities associated with electricity consumption.

6.1 Setup

I use the econometric oligopoly model for the policy simulations. I fix the reference year to 2011. Unless otherwise indicated, the ES requirement is the level in effect in 2011, i.e., 20% relative to the minimum standard. In 2011, the EPA did not require third-party verification for certified models; firms could then certify refrigerator models at no cost, which I also assume here. I use the same choice set as the one used for the supply-side estimation, which also aims to represent the national US refrigerator market for the year 2011 (see Appendix C.1 for more details).

I simulate the demand model for all scenarios with a sample of households taken from the transactions used for the demand estimation. Therefore, households differ with respect to demographic information and the region where they live. The price of electricity faced by each household is the average electricity price at the county level. In one set of simulations, I vary the electricity prices. Finally, I set the rebate level for ES products to zero in all regions. My goal is to focus on the informational dimension of the ES program. Subsidies related to the program are policy add-ons that raise fiscal considerations, which I leave for future research.

My demand model rationalizes observed decisions. It is thus a model of decision utility. I must then make assumptions to recover a measure of consumer welfare based on experienced utility. The gap between decision and experienced utility comes from two sources in my framework. First, there are the costs of collecting and processing energy information, which means that some consumers may make a purchase decision not fully informed. Thus, we must make assumptions regarding what imperfectly informed consumers would experience after their purchase decisions. Given that all consumers will eventually pay for the energy operating costs of their refrigerators,[17] the utility specification of the perfectly-informed consumers ($e = I$) provides the basis to quantify experienced utility. For the welfare calculations, I assume there is no gap between decision and experienced utility for the informed consumers. For the other types, I use the alternative-specific utility at $e = I$ to quantify experienced utility.

The second discrepancy between decision and experienced utility comes from the ES labeling effect. The demand estimates suggest that some consumers have a large willingness to pay for the label, which goes much beyond the value of the energy savings associated with ES-certified products. This label effect could be a bias or something truly experienced by consumers (e.g., warm glow). When quantifying consumer welfare, I will treat it as a bias. I am thus conservative in the benefits that I attribute to the ES program.

In Appendix D.1, I present an expression to quantify consumer welfare accounting for the gap between decision and experienced utility. I derived this expression in Houde (2018) based on Leggett (2002) who first showed how to derive an expression for compensating variation (CV) with imperfect information in a discrete choice framework.[18] Note that the proposed measure does not consider the costs of processing and collecting information to be genuinely experienced. They are considered a psychic cost that does not impact welfare.[19] This assumption is in line with other models of salience and inattention (e.g., Chetty et al. 2009; Farhi and Gabaix 2020).

To compute the externality costs associated with the electricity consumption of refrigerators, I account for the emissions of carbon dioxide (CO_2), sulfur dioxide (SO_2), and nitrous oxide (NO_x). I compute the dollar damages of the externality costs under each scenario by taking the product of the average electricity consumption purchased, the emission factors, and the damage costs per unit of emissions. Table C.3 (Appendix B) presents the emission factors and external damage costs I use. For the welfare calculations, I report two sets of results: one corresponding to the lower end of

[17] Remember that I restrict the estimation sample to homeowners living in a single-family house. Most, if not all, transactions should correspond to households that have to pay for their electricity bills and, thus, ultimately pay the energy operating costs of their appliances.

[18] Allcott 2013; Dubois et al. 2017; Ketcham et al. 2016 also derived similar expressions to conduct behavioral welfare analysis with structural econometric models.

[19] Note, I do not estimate the costs of processing and collecting information.

the estimates of the externality costs and another corresponding to the higher end. These low/high estimates translate into an average external cost of \$0.024/kWh and \$0.079/kWh, respectively.

Finally, simulating the model brings some computational challenges. In a scenario where ES is in effect, the strategy space is discrete-continuous—the existence and uniqueness of a Nash equilibrium are therefore not guaranteed. I solve for the Nash equilibrium using the Gauss-Siedl best-response iteration algorithm, where I randomize the order in which each player plays. I show that the algorithm performs well in Appendix D.2. In scenarios without certification, the equilibrium exists and tends to be unique. In scenarios with certification, multiple equilibria exist, but they predict outcomes close to each other. In particular, although the location of each product in the energy efficiency-price space varies across equilibria, the relative position of each product tends to be the same. In Appendix D.2, I also quantify the uncertainty in the results due to multiple equilibria. The important takeaway is that the main conclusions I present next are robust and quantitatively similar across equilibria.

6.2 Removing the ES Certification

Figure 4 shows the distributions of energy efficiency with and without the ES certification. The dotted histogram is the empirical distribution predicted by the model when the ES requirement is 20% more stringent than the federal minimum standard.[20] The dark histogram represents the predicted distribution without the ES certification and shows that it is optimal for firms to differentiate in the energy dimension. In the absence of the ES certification, firms offer a larger share of products, approximately 15%, that meet the minimum standard. Otherwise, the distribution shifts to the right, i.e., firms increase the efficiency levels of the remaining models well above where the ES requirement was.

Without certification, the energy consumption of the products offered on the market decreases by 11.6 kWh/y, on average, relative to the market with certification (Table 2)—the excess bunching at the ES certification thus crowds out the provision of energy efficiency. Considering the change in demand, removing the ES certification decreases the energy consumption by 12.3 kWh/y, which translates into a decrease in externalities costs ranging from \$3 to \$11 per consumer (Table 2).

Without the ES certification, consumers of all income groups are better off. Consumers benefit for two reasons. First, ES distorts the allocation of energy efficiency by inducing excessive bunching at the requirement. In the absence of ES, firms' energy efficiency offerings better match

[20] Panel (a) of Figure D.2 (Appendix D) compares the distribution observed in the representative choice set for the year 2011 and the simulated distribution. Overall, the model replicates well the excess bunching at the ES requirement.

the distribution of consumer preferences.[21] Second, the ES labeling effect, akin to a strong brand effect, also distorts pricing by enabling firms to set high markups on certified products. These two combined effects imply that removing the certification leads to more economically efficient pricing and product line decisions, benefiting consumers. Figure 5 shows the importance of these effects. I plot the trajectory of each product in the energy efficiency-markup dimensions, where each starting point represents the location of a product with ES, and the endpoint represents the location without ES. We observe a markup reduction for several products, whereas some products are subject to substantial reductions. The reductions are particularly pronounced for products that become less efficient (i.e., the ones moving toward the left in the energy efficiency space). Since firms now offer several products at the lower end of the energy efficiency spectrum, price competition becomes more intense in this region of the product space. As a result, although less energy-efficient, the pricing of these products is more economically efficient, which is also a source of welfare gain.

The impact of ES on firms' profits is small—without the certification, profits decrease by less than \$5/consumer, on average (Table 2). Without certification, firms further differentiate in the energy efficiency dimension. By screening consumers, firms can extract higher markups at the upper end of the energy efficiency spectrum. Figure 5 shows this effect—only products that become more efficient (i.e., move to the right) are subject to higher markups. This figure illustrates the importance of considering firms' strategic behavior and market power. Whereas ES induces excess bunching, in other words, not enough differentiation; removing the certification might do the opposite and induce too much differentiation. As I discuss below, this brings a trade-off in determining the optimal stringency of the ES requirement.

Across income groups, consumers in the middle-income group gain the least from having the certification removed. Relative to the two other income groups, they have a smaller share of consumers relying on the ES certification in the base-case scenario. It thus explains why they are less impacted.

Irrespective of income, informed consumers are all better off without ES. Because they value the full spectrum of energy efficiency, they are the most negatively affected by bunching, especially the crowding out of high-efficiency products. For uninformed consumers, this is the opposite; they tend to be better off when ES is in effect. Inattentive consumers would always be better off in an environment where products are less differentiated. In the extreme scenario where there would be no variation in energy efficiency across products, inattention to energy costs would be, in fact,

[21] Due to imperfect competition, the provision of energy efficiency is, however, still distorted without ES. As in Mussa and Rosen (1978), there should be excess differentiation in the energy dimension.

completely internalized. In essence, this is what ES partially does by reducing the range over which products span in the energy dimension.

In summary, removing the ES certification improves welfare because it leads to more differentiation in the energy efficiency dimension. Although the mass of products that meet the minimum standard increases, the increase in efficiency for other products is significant enough to improve the overall provision of energy efficiency. Consumers benefit from this reallocation of products in the energy dimension.

6.3 Supply-Side Response

How crucial is it to consider firms' strategic response, and by what margins can firms exploit the certification to extract consumer surplus? I provide two sets of simulations to answer those questions. First, I show how the welfare estimates differ under the assumption that firms set energy efficiency levels in a naive manner without certification. In particular, in this counterfactual, I mimic the assumption made by government agencies evaluating the ES program (EPA 2009), which typically assumes that in the absence of the ES certification, all products offered would have the minimum efficiency level prescribed by federal standards. This is a simplistic assumption. In the presence of ES, firms are highly strategic, as shown by their ability to systematically bunch at the ES requirement. We should thus expect a similar level of sophistication without the ES certification. Table 3 shows that ignoring firms' strategic behavior leads to a vast overestimation of the benefits of the ES program.[22] Without ES, we are now facing significant welfare losses: consumers are the worst off, externality costs rise sharply, and the effect on firms' profits is nearly zero. This exercise shows that firms' product differentiation strategies are crucial for determining not only the magnitude but also the direction of the welfare effects. In this context, ignoring firms' strategies leads to an incorrect assessment of the sign of the welfare effects of the policy.

Firms' strategic behaviors are key to determining the welfare effect of the policy. It impacts welfare through two margins: adjustments in quality (i.e., energy efficiency) levels and prices. Quantitatively, which margin is the most significant? I demonstrate that the strategic adjustment in energy efficiency is the most critical. For this analysis, I consider three sets of simulations. First, I hold constant the prices and energy efficiency levels and only remove the certification. This counterfactual captures the pure informational effect of the certification as noted in Houde (2018). Second, I fix the energy efficiency levels and only endogenize pricing strategies, both with and

[22]To implement this counterfactual analysis, I exogenously set the energy efficiency level of each product as follows. Without ES, all products meet the minimum standard. With ES, I use the distribution observed in 2011. I simulate the optimal prices for these predetermined energy efficiency levels for both scenarios.

without ES. Third, I do the opposite—fixing the prices and endogenizing energy efficiency levels. In Table 4, I compare each set of counterfactuals to the main results presented in Table 2. Two important insights emerge. First, in the first and second columns, the changes in the different welfare metrics are small. The pure informational effect of the label is particularly minor. Allowing firms to optimize prices leads to a change in profits close to zero. The changes in externality costs are also essentially zero, which indicates that market shares change little in this scenario. This suggests that when holding energy efficiency levels constant, the optimal equilibrium prices are nearly the same with and without the certification. This is an outcome reminiscent of pure Bertrand competition. When firms compete solely on prices, they cannot extract consumers' willingness to pay associated with the ES label. Conversely, if we allow firms to make optimal adjustments in energy efficiency, this leads to differentiation with and without ES. In the third column, the welfare effects are significant, albeit with different signs compared to the main results. This is the second and most important takeaway: the combined adjustments in quality and price dimensions enable firms to exploit the certification to extract rents. Both margins must be jointly considered to determine the impact of the certification considering strategic behaviors.

6.4 Robustness

Before considering additional policy scenarios, I discuss the robustness of the main results. In particular, I investigate the role of the supply-side response, multiple equilibria, heterogeneity in the cost function, the latent probabilities in the demand model, and long-run dynamics.

6.4.1 Multiple Equilibria

The ES certification creates a discontinuity in the strategy space. As a result, I cannot guarantee either the existence or the uniqueness of a Nash equilibrium. In practice, equilibrium existence is not an issue. I use the Gauss-Siedel best-response iteration algorithm, which converges across most scenarios. However, there are multiple equilibria. In Appendix D.2, I discuss the issue in more detail. The important takeaway is that across different equilibria, the location of the products in the energy efficiency-price space is very consistent. I compute the welfare metrics at the different equilibria and quantify the variation across equilibria; I find that the uncertainty is negligible (see Table D.1). For the policy analysis, the results for each scenario correspond to a particular equilibrium, which I did not select among all the possible local equilibria. All estimates have thus a small degree of uncertainty due to multiple equilibria, but my main conclusions are qualitatively the same across these local equilibria.

6.4.2 Heterogeneous Cost Function

An important parameter to characterize the supply-side is the marginal cost of providing energy efficiency. In practice, we should expect this parameter to vary along different dimensions. For instance, some brands could have a comparative advantage with respect to energy efficiency. The marginal cost could also vary with other attributes such as size or overall design (e.g., side-by-side versus top-freezer refrigerators), which explain most of the variation in refrigerator electricity usage.

Using the pair-matching estimator, I investigate heterogeneity along a few attributes. Table D.2 presents two sets of results. First, I consider heterogeneity induced by the freezer location, which is an important design element for energy efficiency. The estimate for the dummy variable identifying side-by-side and bottom-freezer refrigerators suggests that these designs have a lower marginal cost than top-freezer refrigerators. The estimate is, however, imprecisely estimated. In a second regression, I consider the role of refrigerator size. I found a similar pattern. Larger refrigerators might have a lower marginal cost, but the estimate is also imprecise.

Although heterogeneity in the cost function is not precisely estimated, I perform simulations to show how a richer cost function might influence the results. To illustrate, I focus on heterogeneity with respect to the freezer location. The results are presented in Appendix D.4. Compared to the observed distribution in 2011, the model also does not fit the observed distribution of energy efficiency. The model with cost heterogeneity tends to predict less bunching at the ES requirement and a larger share of highly-efficient products (Figure D.3). Overall, the main conclusions hold, however (Table D.3). The ES tends to crowd out high-efficiency products; ultimately, welfare improves when ES is not in effect.

6.4.3 Impact of Consumer Sophistication

The share of informed/uninformed consumers is an important determinant of welfare. The crowding-out effect occurs because the share of informed consumers is large enough that firms have the incentive to offer highly efficient products. The ES certification should be most beneficial when the share of uninformed consumers dominates.

The latent probabilities $H_{irt}(e)$ in the demand model determine the share of consumers falling into the different types. These probabilities take the form of a multinomial logit where demographic variables and other variables that describe the choice set characterize them. For the demand estimation, the multinomial logit has three options, one for each type ($e = I, ES, U$). To simulate the demand model without ES, I remove the ES type such that the multinomial logit has only two remaining options ($e = I, U$). The variables that enter the latent probabilities determine how the ES type is "reallocated" to the two other types. For instance, in the main policy analysis, the share of informed consumers is 35% with ES and increases to 56% without certification, and the share of uninformed consumers is 36% with ES and increases to 44%.[23] Therefore, two-thirds of ES consumers, which corresponds to 29% of the population in the base-case scenario, are reallocated to the informed type. The other one-third goes to the uninformed type. The estimated coefficients of the latent probabilities dictate this split.

In the following scenarios, I explore how the results vary as a function of the share of informed/uninformed consumers. To do so, I specify the multinomial logit model for the latent probabilities with only two constants, which I then vary exogenously. This approach allows me to generate different shares of consumer types and simulate the model under these scenarios.

Table D.4 presents five different scenarios. The first column is the scenario with the lowest share of informed consumers (9%), and the fifth column has the highest share (67%). The other columns correspond to intermediate scenarios.

Across all five scenarios, removing ES always reduces welfare. As we increase the share of informed consumers, the loss in welfare associated with removing the certification decreases. It is to be expected. Informed consumers do not rely on ES and benefit from having less bunching at the requirement. However, note that none of these scenarios fully capture the crowding-out effect found in the base-case scenario, where removing ES improves welfare. A crucial difference is that in the base-case scenario, the share of ES consumers was much higher compared to the other five scenarios combined with the fact that the share of informed consumers is also significant.[24] As shown in Figure D.4, this explains why bunching is more prevalent and energy efficiency is more differentiated once we remove ES in the base-case scenario. The other scenarios can not replicate these patterns.

Overall, the results show that consumers' abilities to account for energy costs are a crucial determinant of welfare and whether a technology should be subject to ES or not. Given that the ES certification is rolled out on a technology-by-technology basis, the EPA should thus carefully assess the degree of consumers' sophistication before deploying the certification in a specific environment.

[23] Note that these percentages correspond to averages across the whole population of consumers. The latent probabilities vary with demographics and across local markets.

[24] In the first column, the share of ES consumers is about the same as in the base-case scenario. However, the share of informed consumers is much lower, preventing the crowding out of energy efficiency.

6.4.4 Long-Run Dynamics

The oligopoly model endogenizes pricing and one dimension of product line decisions. In the long run, manufacturers, brands, and retailers have, however, several other margins to adjust. For instance, manufacturers bundle other non-energy attributes with ES to create higher-quality products, which may contribute to creating a halo effect associated with ES-certified models (Houde 2022). The fact that the EPA increases the stringency of the ES requirements over time could also be a driver of innovation for manufacturers. In the downstream market, retailers make strategic product placement in stores to increase awareness of the ES program. Brands also use the ES program as a key element of their marketing strategy.[25] In the long run, these different strategies create a strong branding effect associated with ES, and enable firms to exercise market power at various stages of the vertical structure of the market. The present analysis is thus conservative in its representation of market power.

6.5 The Optimal Certification Requirement

The welfare effects depend on various features of the market environment. In the main results, the crowding out of highly efficient products due to the excess bunching at the ES requirement is the leading cause of welfare loss. The regulator could, however, mitigate this effect by setting a more stringent certification requirement.

Figure D.5 (Appendix D) illustrates this and compares the distributions of energy efficiency with and without the certification for different stringency requirements. The two distributions coincide for more stringent requirements, and there is no crowding out. The welfare estimates for the different stringency requirements suggest that setting the stringency between 30% and 40%, which corresponds to about a doubling of the stringency that was in effect in 2011, would have made ES welfare improving.

Another margin that policymakers can use in the design of certification is its level of coarseness. The ES certification, for instance, introduced a higher-tier program called the ”Most Efficient” in 2011 to further differentiate certified products. In other markets, multi-tier certification schemes are the norm. In Europe, for example, cars, appliances, and houses are labeled with an A to F scale. How does coarseness impact welfare? To explore this question, I compare a binary certification scheme to a two-tier scheme. This analysis, however, requires additional assumptions about how consumers integrate such information and value different tiers. Consumers can now be categorized

[25] As part of the ES program, the EPA has created the “Partner of the Year” Excellence Award. Kenmore is one brand that has made a sustained effort to get this award to market itself as a sustainable firm. The program also rewards retailers.

into four types. As before, they can be entirely inattentive to energy efficiency-related information or act rationally by paying attention to the continuous measure of energy-related information while ignoring the redundant ES certification. Alternatively, they can focus on either the lower tier or the higher tier of the certification. I will assume consumers sort into these two latter categories based on the same attention allocation parameters estimated in the demand model, where the share of certified models offered in each tier category will determine the shares of consumers attending each certification level. If consumers pay attention to one of the tiers, they will value certified products based on their expected savings, which they compute using rational expectations based on their knowledge of the tier threshold and overall distribution of energy costs. They also value certified products due to a pure label effect, represented by the parameter τ^{ES} in the demand model. I will consider three cases to account for the label effect in the presence of multiple tiers. First, I assume an equal splitting of the label effect estimated in the binary scheme between the two tiers, i.e., $\tau^{ES}/2$ at each tier. Second, I assume that the label effect entirely transposes to the higher tier, making the label effect zero at the lower tier. Third, I assume that the label effect is the same for both tiers and is not diluted, i.e., it is equal to τ^{ES} at each tier.

To illustrate the role of multiple tiers, I compare a scenario where the certification is binary, and the certification requirement is set close to the optimum at 30% more stringent than the federal minimum energy efficiency standard with a scenario where the scheme has two tiers, where the lower tier is at 15%, and the highest tier is at 30%. Table 5 presents the results. Compared to the binary case, the multiple-tier scheme leads to overall welfare loss,[26] regardless of how I model the label effect. To gain further insights into the source of the welfare losses, Panels a) and b) of Figure 6 present the corresponding energy efficiency distributions under the different scenarios. The multiple-tier scheme smooths the distribution of energy efficiency compared to the binary case, which impacts welfare via three channels. First, although it limits allocative distortion by offering consumers more options along the energy efficiency dimension, it discourages the adoption of more energy-efficient options. The binary scheme leads to a larger reduction in overall externality costs, unlike the two-tier scheme. Second, the label effect, which is treated as a source of bias in the welfare calculations, could be exacerbated by adding multiple tiers. Note that this latter effect is highly dependent on how we interpret the high willingness to pay for the label itself. Third, the two-tier scheme reduces firms' ability to charge high markups. Firms are better off under a binary scheme compared to the two-tier scheme. Adding a tier increases the mass of products at the center of the energy efficiency distribution, which reduces differentiation among certified products.

[26] The results present the welfare effect of removing the certification. A positive welfare estimate means that the market without certification is preferable. Alternatively, a negative value for the change in profits means that firms are better off under the certification.

This, in turn, limits firms' ability to charge higher markups on certified products, resulting in lower profits. This can be clearly seen on Panel c) and d) of Figure 6, which shows the movement of each product in the energy efficiency-markup dimensions. Under the binary scheme, it is much easier for firms to charge large markups on a subset of certified products.

These mechanisms will be transposed to certification schemes with even more tiers. Additional tiers will further dilute firms' ability to exploit certification to extract market power. Ultimately, with a large number of tiers, the equilibrium will converge to the scenario where there is no certification and energy efficiency information is approximately continuous. Firms will thus revert to screening consumers who are completely inattentive and consumers who pay attention to the continuous but more complex measure of energy efficiency.

6.6 Interactions with Electricity Prices

A distinctive feature of the US electricity market is that electricity prices vary significantly across regions. For instance, during my sample period, the 5% percentile in average county electricity price is 0.08 $/kWh, while the 95% percentile is 0.18 $/kWh. This variation is attributable to various institutional features, such as the presence of regional environmental policies, liberalization of electricity markets by state regulators, and proximity to coal and natural gas reserves, to name a few (Zivin et al. 2014). One implication of these heterogeneous electricity prices is that the benefits of policies aimed at reducing energy demand can change widely across regions. Federal demand-side energy policies, however, are typically one-size-fits-all and do not account for such variation. This can lead to losses in economic efficiency and highly heterogeneous distributional impacts.

In the context of the ES certification, variation in local electricity prices implies that the private net benefits of purchasing an ES-certified product would depend on the region where a household lives. In regions with low electricity prices, the reduction in lifetime energy costs associated with ES may be too low to compensate for higher retail prices, and vice versa when electricity prices are high.

In Appendix D, I provide additional results where I decompose the change in consumer surplus of the main results for households living in regions of the United States with low (less than 0.11 $/kWh), medium (0.11 to 0.16 $/kWh), and high (more than 0.16 $/kWh) electricity prices (Table D.5). Across these regions, removing the ES certification increases the consumer surplus for all households across income groups. Households are better off without ES, especially in regions with high electricity prices.

One important related question is: how would the welfare impacts of ES change if electricity prices were reformed? As pointed out by Borenstein and Bushnell (2022), electricity prices are widely inefficient across the US for several reasons. First, the negative environmental externalities associated with electricity production are not systematically internalized in electricity prices. Second, the fixed costs of electricity production are often recovered as volumetric charges, which results in high average prices. Market power is also an issue. These have countervailing effects. Ultimately, the final price residential consumers pay is either too high or too low relative to the social marginal cost.[27]

To illustrate the interaction between an information-based policy such as ES and a price-based approach to manage energy demand, I use Borenstein and Bushnell (2022)'s prices in my welfare analysis. Because their data do not precisely cover my sample period, I first simulate the model with their average prices, which are constructed using an approach similar to mine. I then consider that all households pay electricity prices corresponding to the social marginal costs in their region (viz., the private marginal cost of electricity generation plus the marginal environmental costs).

Table D.6 presents the results. The first column corresponds to my base-case scenario. The second column presents the results for the average prices constructed by Borenstein and Bushnell (2022), which are consistent with the base case. Overall, ES is crowding out energy efficiency; removing the certification improves welfare. The third column presents the results when consumers pay electricity prices set at their social marginal costs. The sign of the welfare changes is now reversed. The ES certification is now welfare-improving. The intuition is simple. Under this scenario, electricity prices are now lower, on average. Consumers thus have a lower demand for energy-efficient products. In turn, firms offer less efficient products (Figure D.6). Without inattentive consumers, the energy efficiency offered under social marginal costs would be economically efficient. Inattention is, however, present and consists of an additional market failure. Firms' product offering in the energy efficiency dimension is thus socially inefficient even if electricity prices are set at their social marginal cost. In this context, there is thus a role for an information-based policy that would correct consumers' inattention. This is what the ES program does, albeit imperfectly due to the coarseness of the information it provides.

[27] Borenstein and Bushnell (2022) computed the difference between the average electricity price and the social marginal price for all US zip codes during the 2014-2016 period and found that average prices were too high compared to the socially optimal level for several US regions.

7 Conclusion

This paper develops a framework to study the welfare effects and incidence of coarse certification programs. The framework accounts for the strategic behavior of firms and the various behavioral mechanisms by which such certification can influence consumers. In particular, consumers trade off the more accurate but difficult-to-collect and process information for the coarse but simple and salient signal provided by a certification. I show that this trade-off can lead to an unintended consequence—a coarse certification may crowd out the provision of the hard-to-assess dimension of quality and induce excess bunching at the certification requirement.

I apply the framework to the ENERGY STAR program, one of the most important US federal policies to manage energy demand. Using the refrigerator market as a case study, I show that in the absence of the ES certification, products will be more differentiated in the energy efficiency dimension. A large share of products would meet the minimum energy efficiency standard, and another share would have high-efficiency levels that exceed the ES certification requirement that was in effect. Overall, the average energy consumption of the products offered on the market decreases without certification—ES thus crowds out the provision of energy efficiency. Consumers are also better off without certification. The effect of the certification on firms' profits is small and heterogeneous.

The policy analysis offers a cautionary tale on how certification, and more generally, nudges and information-based policies, should be used, especially when addressing environmental externalities. Historically, the US Government has managed the ES program like a marketing program where a strong branding effect has been sought and deemed a successful metric. I show that consumers' willingness to pay for the ES label favors the adoption of certified products and induces firms to offer more of these products. However, this bunching at the ES requirement may not necessarily translate into improvement in energy efficiency levels and welfare.

References

Akerlof, G. A.: 1970, The market for 'lemons': Quality uncertainty and the market mechanism, *The Quarterly Journal of Economics* **84**(3), 488–500.

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

Allcott, H. and Sweeney, R. L.: 2016, The role of sales agents in information disclosure: Evidence from a field experiment, *Management Science* **63**(1), 21–39.

Amano, T. and Ohashi, H.: 2018, Ratcheting, competition, and the diffusion of technological change: the case of televisions under an energy efficiency program, *Columbia Business School Research Paper* (18-71).

Ashenfelter, O. C., Hosken, D. S. and Weinberg, M. C.: 2013, The price effects of a large merger of manufacturers: A case study of Maytag-Whirlpool, *American Economic Journal: Economic Policy* **5**(1), 239–61.

Bagwell, K.: 2007, The economic analysis of advertising, Vol. 3 of *Handbook of Industrial Organization*, Elsevier, pp. 1701–1844.

Barahona, N., Otero, C. and Otero, S.: 2023, Equilibrium effects of food labeling policies, *Econometrica* **91**(3), 839–868.

Berto Villas-Boas, S.: 2007, Vertical relationships between manufacturers and retailers: Inference with limited data, *The Review of Economic Studies* **74**(2), 625–652.

Boatwright, P., Kalra, A. and Zhang, W.: 2008, Research note—should consumers use the halo to form product evaluations?, *Management Science* **54**, 217–223.

Bolton, P. and Dewatripont, M.: 2005, *Contract Theory*, MIT press.

Bonnet, C. and Dubois, P.: 2010, Inference on vertical contracts between manufacturers and retailers allowing for nonlinear pricing and resale price maintenance, *The RAND Journal of Economics* **41**(1), 139–164.

Bonroy, O. and Constantatos, C.: 2014, On the economics of labels: How their introduction affects the functioning of markets and the welfare of all participants, *American Journal of Agricultural Economics* **97**(1), 239–259.

Borenstein, S. and Bushnell, J. B.: 2022, Do two electricity pricing wrongs make a right? Cost recovery, externalities, and efficiency, *American Economic Journal: Economic Policy* **14**(4), 80–110.

Chetty, R., Looney, A. and Kroft, K.: 2009, Salience and taxation: Theory and evidence, *American Economic Review* **99**(4), 1145–77.

DellaVigna, S.: 2018, Structural behavioral economics, *Handbook of Behavioral Economics: Applications and Foundations 1*, Vol. 1, Elsevier, pp. 613–723.

Dranove, D. and Jin, G. Z.: 2010, Quality disclosure and certification: Theory and practice, *Journal of Economic Literature* **48**(4), 935–963.

Dubois, P., Griffith, R. and OConnell, M.: 2017, The effects of banning advertising in junk food markets, *The Review of Economic Studies* **85**(1), 396–436.

Eichholtz, P., Kok, N. and Quigley, J. M.: 2010, Doing well by doing good? Green office buildings, *American Economic Review* **100**(5), 2492–2509.

EPA: 2009, Annual Report - ENERGY STAR, *Technical report*, U.S. EPA, Office of Air and Radiation, Climate Protection Partnerships Division.

Farhi, E. and Gabaix, X.: 2020, Optimal taxation with behavioral agents, *American Economic Review* **110**(1), 298–336.

Fischer, C. and Lyon, T. P.: 2014, Competing environmental labels, *Journal of Economics & Management Strategy* **23**(3), 692–716.

Greenstone, M., Kopits, E. and Wolverton, A.: 2013, Developing a social cost of carbon for us regulatory analysis: A methodology and interpretation, *Review of environmental economics and policy* .

Harbaugh, R., Maxwell, J. W. and Roussillon, B.: 2011, Label confusion: The groucho effect of uncertain standards, *Management Science* **57**(9), 1512–1527.

Hausman, J. A.: 1996, Valuation of new goods under perfect and imperfect competition, *The Economics of New Goods*, NBER Chapters, National Bureau of Economic Research, Inc, pp. 207–248.

Heyes, A. and Martin, S.: 2016, Social labeling by competing NGOs: A model with multiple issues and entry, *Management Science* **63**(6), 1800–1813.

Holland, S. P., Hughes, J. E. and Knittel, C. R.: 2009, Greenhouse gas reductions under low carbon fuel standards?, *American Economic Journal: Economic Policy* **1**(1), 106–146.

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

Houde, S.: 2022, Bunching with the stars: How firms respond to environmental certification, *Management Science* **68**, 5569–5590.

Houde, S. and Aldy, J. E.: 2017, Consumers' response to state energy efficient appliance rebate programs, *American Economic Journal: Economic Policy* **9**(4), 227–255.

Houde, S. and Myers, E.: 2021, Are consumers attentive to local energy costs? evidence from the appliance market, *Journal of Public Economics* **201**, 104480.

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

Jacobsen, M. R.: 2013, Evaluating U.S. fuel economy standards in a model with producer and household heterogeneity, *American Economic Journal: Economic Policy* **5**(2), 148–187.

Kahneman, D., Wakker, P. P. and Sarin, R.: 1997, Back to bentham? explorations of experienced utility*, *The Quarterly Journal of Economics* **112**(2), 375–406.

Keane, M., Ketcham, J., Kuminoff, N. and Neal, T.: 2021, Evaluating consumers' choices of medicare Part D plans: A study in behavioral welfare economics, *Journal of Econometrics* **222**(1), 107–140.

Ketcham, J. D., Kuminoff, N. V. and Powers, C. A.: 2016, 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.

Klier, T. and Linn, J.: 2012, New-vehicle characteristics and the cost of the corporate average fuel economy standard, *The RAND Journal of Economics* **43**(1), 186–213.

Knittel, C. R.: 2011, Automobiles on steroids: Product attribute trade-offs and technological progress in the automobile sector, *American Economic Review* **101**(7), 3368–99.

Kotchen, M. J.: 2006, Green markets and private provision of public goods, *Journal of Political Economy* **114**(4), 816–834.

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

Mason, C. F.: 2011, Eco-labeling and market equilibria with noisy certification tests, *Environmental and Resource Economics* **48**(4), 537–560.

Mullainathan, S., Schwartzstein, J. and Shleifer, A.: 2008, Coarse thinking and persuasion, *The Quarterly Journal of Economics* **123**(2), 577–619.

Muller, N. Z. and Mendelsohn, R.: 2012, Efficient pollution regulation: Getting the prices right: Corrigendum (mortality rate update), *American Economic Review* **102**(1).

Mussa, M. and Rosen, S.: 1978, Monopoly and product quality, *Journal of Economic Theory* **18**(2), 301–317.

Reynaert, M. and Sallee, J. M.: 2021, Who benefits when firms game corrective policies?, *American Economic Journal: Economic Policy* **13**(1), 372–412.

Rysman, M., Simcoe, T. and Wang, Y.: 2020, Differentiation strategies in the adoption of environmental standards: Leed from 2000 to 2014, *Management Science* **66**(9), 4173–4192.

Sallee, J. M.: 2014, Rational inattention and energy efficiency, *The Journal of Law and Economics* **57**(3), 781–820.

Sims, C. A.: 2003, Implications of rational inattention, *Journal of Monetary Economics* **50**(3), 665–690.

Small, K. A. and Rosen, H. S.: 1981, Applied welfare economics with discrete choice models, *Econometrica* **49**(1), 105–130.

Spurlock, C. A. and Fujita, K. S.: 2022, Equity implications of market structure and appliance energy efficiency regulation, *Energy Policy* **165**, 112943.

Stahl, K. and Strausz, R.: 2017, Certification and market transparency, *The Review of Economic Studies* **84**(4), 1842–1868.

Stole, L. A.: 1995, Nonlinear pricing and oligopoly, *Journal of Economics & Management Strategy* **4**(4), 529–562.

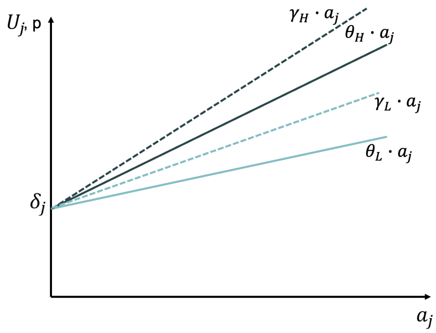
Walls, M., Gerarden, T., Palmer, K. and Bak, X. F.: 2017, Is energy efficiency capitalized into home prices? evidence from three us cities, *Journal of Environmental Economics and Management* **82**, 104–124.

Ward, D. O., Clark, C. D., Jensen, K. L., Yen, S. T. and Russell, C. S.: 2011, Factors influencing willingness-to-pay for the ENERGY STAR label, *Energy Policy* **39**(3), 1450–1458.

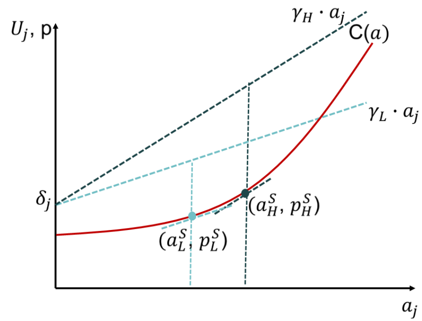
Whitefoot, K. S., Fowlie, M. L. and Skerlos, S. J.: 2017, Compliance by design: Influence of acceleration trade-offs on CO2 emissions and costs of fuel economy and greenhouse gas regulations, *Environmental Science & Technology* **51**(18), 10307–10315.

Zivin, J. S. G., Kotchen, M. J. and Mansur, E. T.: 2014, Spatial and temporal heterogeneity of marginal emissions: Implications for electric cars and other electricity-shifting policies, *Journal of Economic Behavior & Organization* **107**, 248–268.

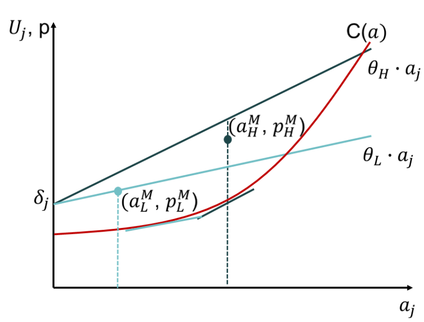
8 Figures and Tables



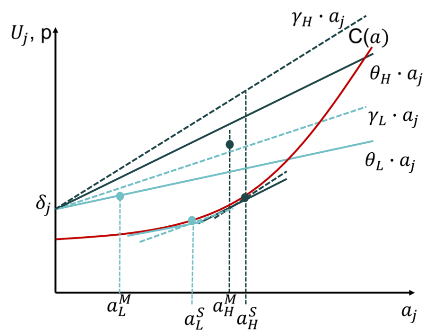
(a) Consumer Utility: Experienced vs. Decision Utility



(b) Social Outcome



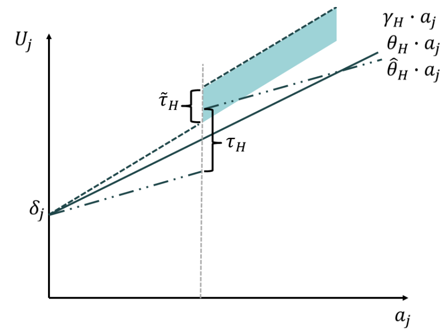
(c) Monopoly Outcome



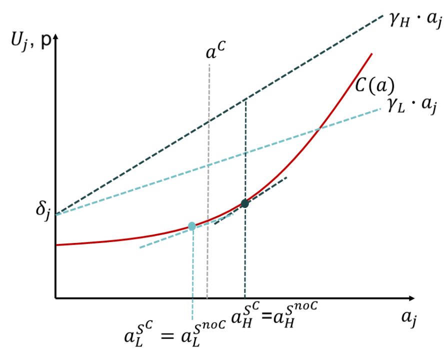
(d) Monopoly vs. Social

Figure 1: Equilibrium Outcomes without Certification

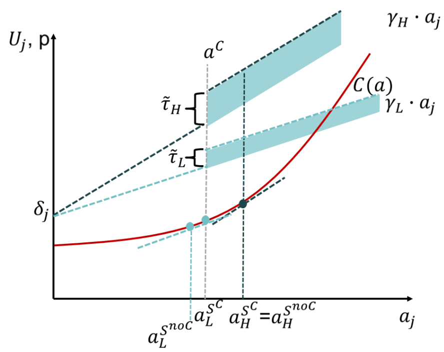
Notes: Panel (a) shows consumer utility for good j without a coarse certification. Experienced utility is depicted by the dashed lines, and decision utility is depicted by the solid lines. In Panel (b), the social outcome corresponds to the points where the experienced marginal valuations equal the marginal cost. The monopoly outcome, Panel (c), is the solution of the standard screening problem of Mussa and Rosen (1978), except that the monopolist considers consumers' decision utility. The incentive compatibility constraint of the high type explains why the monopolist cannot completely extract the consumer surplus of the high type. Relative to the social outcome, the monopolist under-provides quality to the low and high types, Panel (d).



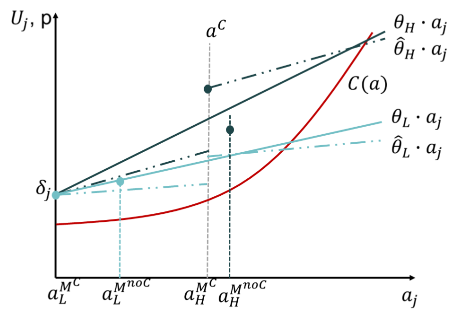
(a) Consumer Utility: High Type Only



(b) Social Outcome: $\tilde{\tau}$ is a Bias



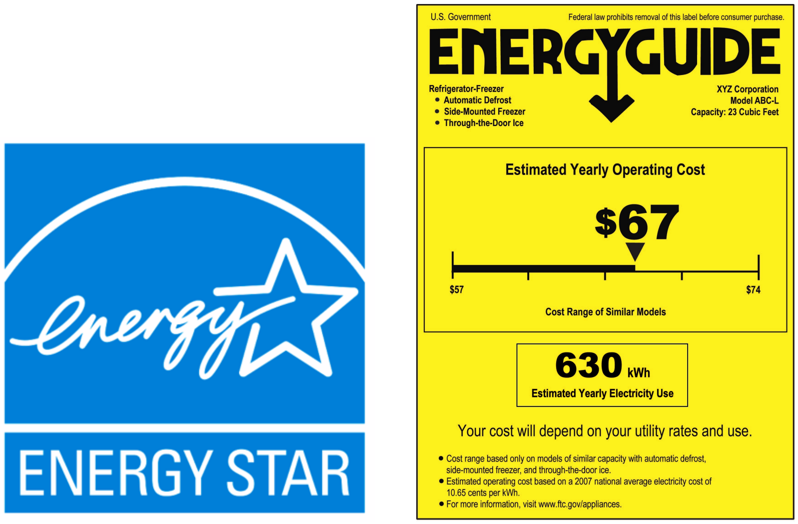
(c) Social Outcome: $\tilde{\tau}$ is Preferences



(d) Monopoly Outcome with Crowding-Out

Figure 2: Equilibrium Outcomes with and without a Coarse Certification

Notes: Panel (a) illustrates how the consumer's utility (high-type only) for good j changes with a coarse certification. The dotted line is the experienced utility. The shaded area represents the gain in utility if the persuasion effect, $\tilde{\tau}$, is considered experienced utility. The certification creates a discontinuous jump in the valuation of a at the certification requirement a^C. Decision utilities without and with certification are shown by the plain and irregular dotted lines, respectively. With certification, the decision utility has a lower marginal valuation of a but a bigger discontinuity at the certification requirement. The social outcomes with certification, denoted with the superscript S^C, and without certification, denoted with the superscript S^{noC}, are shown on panels (b) and (c). Panel (b) represents the case where $\tilde{\tau}$ is a bias. Panel c represents the case where $\tilde{\tau}$ corresponds to preferences. The monopolist outcomes with and without certification are shown on Panel (d) for a case when bunching at the requirement a^C is optimal and crowding out of low and high quality occurs.



(a) ES Label (b) EnergyGuide Label

Figure 3: ENERGY STAR (ES) and EnergyGuide Labels

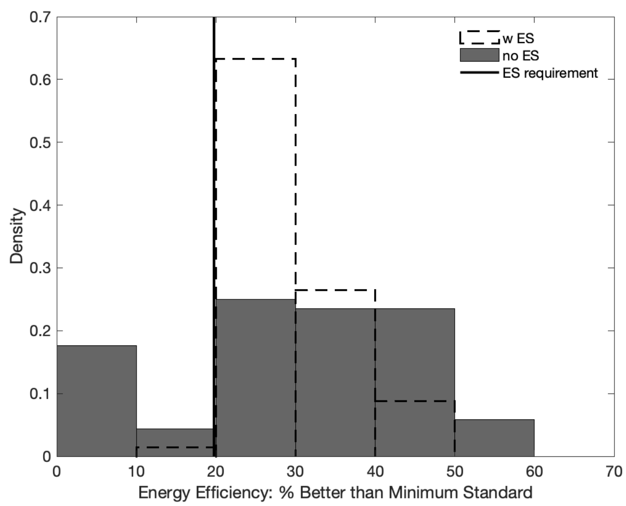


Figure 4: Distributions of Energy Efficiency with and without Certification
Notes: The figure presents the simulated distributions with ES (dotted line) and without ES (dark histogram). The ES requirement is set at 20% more stringent than the federal minimum energy efficiency standard. Comparing the two distributions, we find that the certification crowds out high-efficiency products and removes many products at the lowest efficiency level.

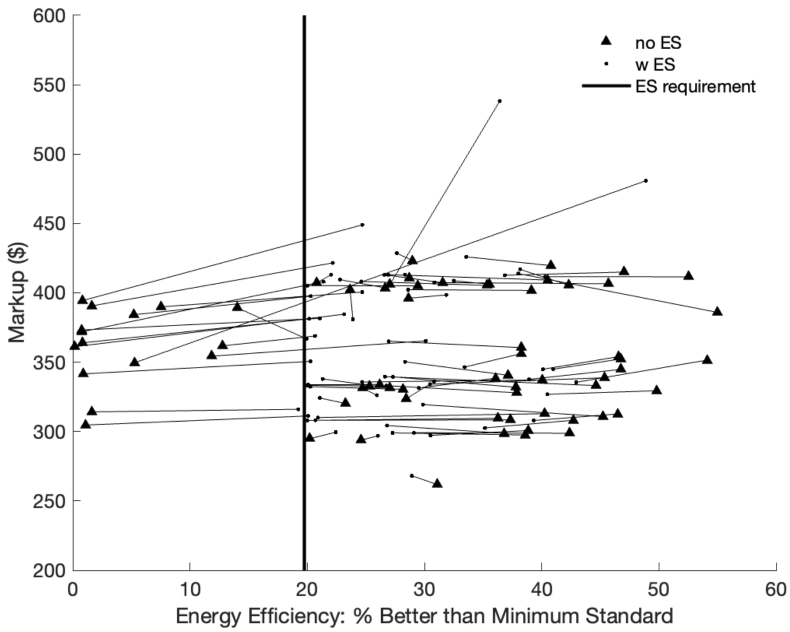
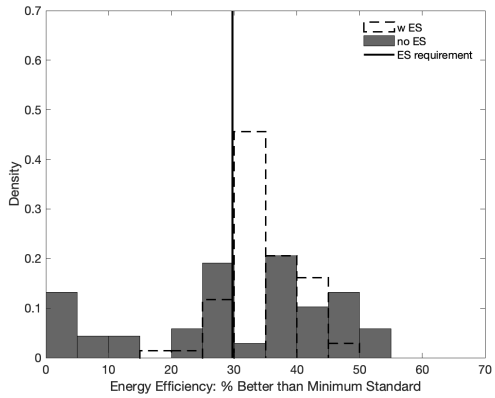
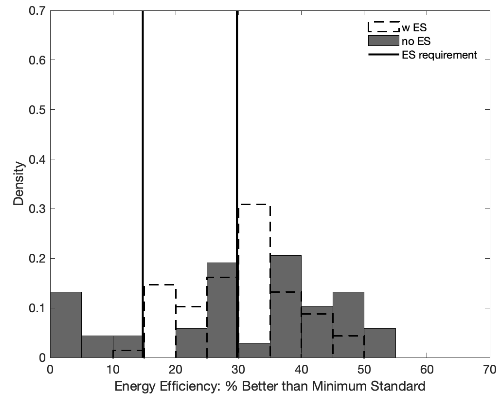


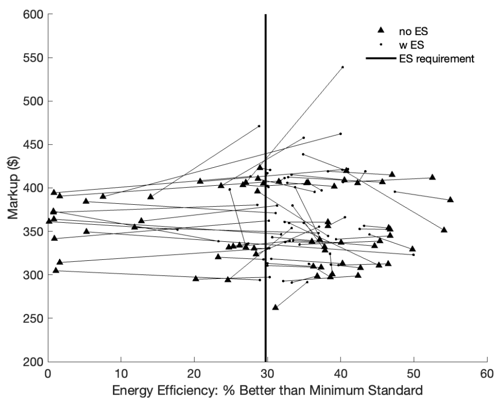
Figure 5: Changes in Energy Efficiency and Markups
Notes: Each marker represents the location of a product in the energy efficiency dimension relative to markups. Triangles represent the scenario without certification, and dots represent the scenario with certification. Arrows depict the trajectory of each product.



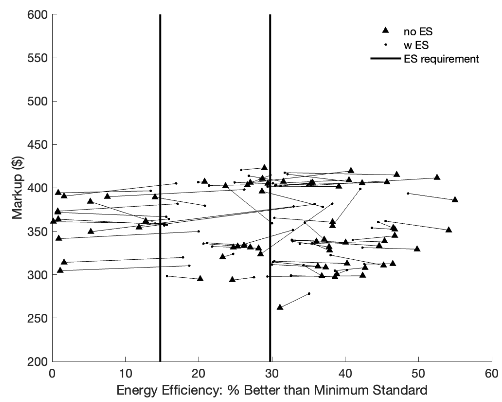
(a) Binary ES



(b) Two-tier ES



(c) Binary ES



(d) Two-tier ES

Figure 6: Energy Efficiency and Markups: Binary versus Two-Tier Certification

Table 1: Paired Refrigerator Models: Summary Statistics and Estimation Results

	Paired Models	All Models
Summary Statistics:		
# 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
Estimation Marginal Cost Providing EE (ϕ):		
$ln(price_{j,r}^{manuf}) = \alpha + \gamma_{j,j'} + Y_j \times Brand_j + \phi Efficiency_j + \epsilon_{j,r}$,		
Pair FE only	182.6	
	(55.5)	
Pair FE & Year-Brand FE	191.1	
	(67.5)	
Hedonic Regression	220.1	
	(97.5)	

Notes: The sample used to identify the identical pairs of refrigerator models contains all models offered by the retailer during the period 1998-2011. Only a subset of those models was used in the demand estimation. The summary statistics show that the paired refrigerator models are smaller, cheaper, and more efficient than the full sample. The parameter ϕ is estimated with three specifications. The hedonic regression does not contain a paired fixed effect but controls for the attributes used to identify the identical pairs: brand, size, door style, door material (stainless or not), ice-maker option, defrost technology, air filtration system, color, and door handle type. Standard errors are in parentheses.

Table 2: The Effects of Removing the ES Certification

Panel A: Change in Attributes				
Δ Sales weighted kWh/y		-12.3		
Δ Offered kWh/y		-11.7		
Δ Sales weighted price ($)		7.7		
Δ Offered price ($)		-3.7		

Panel B: Change in Welfare by Income Group	Income <$50,000	Income ≥$50,000 & <$100,000	Income ≥$100,000	All
Δ CV	18.8	7.0	15.0	12.4
Δ CV, e=I	21.4	16.7	24.0	20.0
Δ CV, e=U	1.4	-10.3	-35.9	-13.9
Δ Externalities-Low	-8.3	-1.1	-2.3	-3.4
Δ Externalities-High	-27.2	-3.6	-7.7	-11.3
Δ Profits				-5.2
Δ Social Welfare-Low				10.7
Δ Social Welfare-High				18.6

Notes: The table reports the difference between a scenario without ES and a scenario with ES. The counterfactual scenario is the market without ES. A negative sign implies that removing ES leads to a decrease in a particular metric relative to a market with ES. All figures are in dollars per consumer. The term CV refers to the compensating variation and quantifies the average change in consumer welfare across all consumer types. The compensating variation for the informed type ($e = I$) and uninformed type ($e = U$) are also presented. The negative externalities associated with electricity consumption are evaluated for two scenarios: 'Low' refers to the lower range of the damage estimates, and 'High' refers to the higher range of the estimates.

Table 3: The Effects of Removing the ES Certification
with Naive Adjustment to Energy Efficiency

Panel A: Change in Attributes				
Δ Sales weighted kWh/y		116.8		
Δ Offered kWh/y		117.3		
Δ Sales weighted price ($)		-44.8		
Δ Offered price ($)		-54.6		

Panel B: Change in Welfare by Income Group	Income <$50,000	Income ≥$50,000 & <$100,000	Income ≥$100,000	All
Δ CV	-78.7	-80.1	-142.4	-96.5
Δ CV, e=I	-79.4	-81.5	-149.2	-99.2
Δ CV, e=U	-87.5	-88.0	-156.8	-106.4
Δ Externalities-Low	29.3	33.9	34.3	32.7
Δ Externalities-High	96.6	111.8	112.8	107.8
Δ Profits				0.0
Δ Social Welfare-Low				-129.3
Δ Social Welfare-High				-204.4

Notes: The table reports the difference between a scenario without ES and a scenario with ES. In the scenario without ES, the energy efficiency levels of all products are set at the minimum standard. With ES, they are set at the observed levels in 2011. Only prices are endogenized. The counterfactual scenario is the market without ES. A negative sign implies that removing ES leads to a decrease in a particular metric relative to a market with ES. All figures are in dollars per consumer. The term CV refers to the compensating variation and quantifies the average change in consumer welfare across all consumer types. The compensating variation for the informed type ($e = I$) and uninformed type ($e = U$) are also presented. The negative externalities associated with electricity consumption are evaluated for two scenarios: 'Low' refers to the lower range of the damage estimates, and 'High' refers to the higher range of the estimates.

Table 4: The Effects of Removing the ES Certification
with Partial Strategic Adjustments

	Only Label Removed	Label Removed Only Price Optimized	Label Removed Only kWh Optimized	Label Removed Price and kWh Optimized
Δ CV	6.9	5.5	-13.0	12.4
Δ CV, e=I	0.0	0.1	-5.1	20.0
Δ CV, e=U	0.0	0.1	-33.8	-13.9
Δ Externalities-Low	0.5	0.0	2.4	-3.4
Δ Externalities-High	1.5	0.0	7.9	-11.3
Δ Profits	-7.3	-0.6	4.5	-5.2
Δ Social Welfare-Low	-0.9	4.9	-10.9	10.7
Δ Social Welfare-High	-1.9	4.9	-16.5	18.6

Notes: The table reports the difference between a scenario without ES and a scenario with ES. In the first column, the energy efficiency and prices are fixed at the observed levels and only the label is removed. In the second column, only prices are optimized with and without the ES label. In the third column, only energy efficiency levels (kWh) are optimized with and without the ES label. The last column presents the main results where both prices and energy efficiency levels are optimized. As in Table 2, a negative sign implies that removing ES leads to a decrease in a particular metric relative to a market with ES. All figures are in dollars per consumer.

Table 5: The Effects of Removing the ES Certification
Binary vs Two-Tier Schemes

	Binary	Two-tier		
	$ES = 30\%$	$ES_{low} = 15\%$, $ES_{high} = 30\%$		
Label Effect:	τ^{ES}	$[\tau^{ES}/2, \tau^{ES}/2]$	$[0, \tau^{ES}]$	$[\tau^{ES}, \tau^{ES}]$
Δ CV	-2.5	6.6	9.1	14.4
Δ CV, e=I	3.7	9.4	9.4	9.4
Δ CV, e=U	-30.3	-21.8	-21.8	-21.8
Δ Externality-Low	7.6	-1.5	-1.1	-1.4
Δ Externality-High	24.9	-5.0	-3.6	-4.7
Δ Profits	-7.1	-1.0	-1.0	-1.0
Δ Social Welfare-Low	-17.1	7.1	9.2	14.8
Δ Social Welfare-High	-34.5	10.6	11.6	18.1

Notes: The table reports the difference between a scenario without ES and a scenario with ES. The first column reports results for a binary certification with a requirement set at 30% below the minimum federal standard. The other columns report results for a two-tier scheme where the low tier has a 15% requirement, and the higher tier has a 30% requirement. The label effect for the binary scheme is equally split $[\tau^{ES}/2, \tau^{ES}/2]$, only allocated to the highest tier $[0, \tau^{ES}]$, or the same for both tiers $[\tau^{ES}, \tau^{ES}]$, respectively. As in Table 2, a negative sign implies that removing ES leads to a decrease in a particular metric relative to a market with ES. All figures are in dollars per consumer.

A Theory: Proofs and Additional Results

Social Outcome The social planner aims to maximize the total surplus and considers the utility experienced by the consumers. He will then set the level of a for each consumer type i such that the marginal valuation of a experienced by each consumer type equals the marginal cost: $\gamma_i = C'(a_i)$.

Monopolist Outcome The monopolist has prior beliefs about θ_L and θ_H, but not the underlying experienced utility. She will then base her optimal strategy based on her beliefs of θ_L and θ_H.

Formally, the problem of the monopolist choosing the level of attribute a and price p becomes the standard monopolist screening problem of Mussa and Rosen (1978):

$$\max_{a_H, a_L, p_H, p_L} \pi \cdot (p_L - C(a_L)) + (1-\pi) \cdot (p_H - C(a_H))$$

s.t.

$$\text{IRH: } \delta + \theta_H a_H - p_H \geq 0$$

$$\text{IRL: } \delta + \theta_L a_L - p_L \geq 0$$

$$\text{ICH: } \theta_H a_H - p_H \geq \theta_H a_L - p_L$$

$$\text{ICL: } \theta_L a_L - p_L \geq \theta_L a_H - p_H$$

The following lemma and proof follow closely Bolton and Dewatripont (2005).

Lemma 1. *At an interior solution:*

- *ICL and IRH are non-binding; and*
- *ICH and IRL are binding.*

Proof. **Step I.** If $\theta_H > \theta_L$, ICH and ICL cannot be both binding in a separating equilibrium ($a_H \neq a_L$, $p_H \neq p_L$).

If ICH and ICL are both binding and $a_H \neq a_L$, this implies $\theta_H = \theta_L$, a contradiction.

Step II. ICi and IRi, $i = \{L, H\}$, cannot be both non-binding in equilibrium.

If both constraints ICi and IRi are non-binding, the firm can increase its profit by slightly increasing p_i, a contradiction.

Step III. ICL is non-binding.

If ICL is binding, ICL and IRL implies that IRH is non-binding:

$$0 \leq \delta + \theta_L a_L - p_L = \delta + \theta_L a_H - p_H \leq \delta + \theta_H a_H - p_H \tag{9}$$

IRH non-binding implies that ICH should bind (by Step II). ICL and ICH are then both binding in equilibrium, a contradiction (by Step I).

Step IV. IRH is non-binding

Step III implies that IRL is binding (by Step II). If IRL is binding, ICH and IRL implies that IRH is non-binding:

$$0 = \delta + \theta_L a_L - p_L < \delta + \theta_H a_L - p_L \leq \delta + \theta_H a_H - p_H \tag{10}$$

Step V. IRL and ICH is binding

Steps III and IV together with Step II, respectively imply that IRL and ICH are binding.

If $\theta_H > \theta_L$, the single crossing condition holds. This ensures that the incentive compatibility constraint of the low type (ICL) is non-binding at the optimum. Moreover, if ICL is not binding at the optimum, the individual rationality constraint of the low type (IRL) must be binding, otherwise the firm could increase profit by slightly increasing the price p_L. By a similar argument, the incentive compatibility constraint of the high type (ICH) must be binding at the optimal, i.e., the consumer with a high valuation of a may have an incentive to purchase the technology offered to the consumer with a low valuation of a. The firm must then distort the prices and attribute a to ensure that it is not optimal for the high type consumer to purchase the technology with $a = a_L$.

Using the fact that IRL and ICH are binding, we can solve for prices as a function of the attribute levels. The relaxed form of the monopolist's problem is given by

$$\max_{a_H, a_L} \pi \cdot (\delta + \theta_L a_L - C(a_L)) + (1-\pi) \cdot (\delta + \theta_H (a_H - a_L) + \theta_L a_L - C(a_H))$$

The first order conditions yield:

$$\theta_H = C'(a_H^*)$$
$$\frac{\theta_L - (1-\pi)\theta_H}{\pi} = C'(a_L^*),$$

and the optimal prices are given by:

$$p_L^* = \delta + \theta_L a_L^*$$
$$p_H^* = \theta_H (a_H^* - a_L^*) + \delta + \theta_L a_L^*$$

Proposition 1. *Relative to the social optimum, the monopolist under-provides quality to the low and high types, and extracts all the consumer surplus from the low type.*

Proof. As in the standard screening problem, quality is under-provided to the low type because the following inequality must hold at a separating equilibrium:

$$\theta_L > \frac{\theta_L - (1-\pi)\theta_H}{\pi}$$

Given that $\gamma_L > \theta_L$, and the cost is increasing in a, i.e., $C'(\cdot) > 0$, there is further distortion in the level of a_L provided by the monopolist.

For the high type, given that $\theta_H < \gamma_H$ and $C'(\cdot) > 0$, the value of a that solves $C'(a_H) = \gamma_H$ is always larger than the value of a that solves $C'(a_H) = \theta_H$. Quality is thus also under-provided to the high type.

Proposition 2. *Define* π_P^**, the profits obtained under a pooling equilibrium where both products bunch at the certification* a^C*, and* π_S^**, the profits obtained under a separating equilibrium, where* $a_L^* \neq a_H^*$ *and* $p_L^* \neq p_H^*$*. If* $a_L^* < a^C$ *and* $\tau_L = \tau_H = \tau$*, we have:*

$$\frac{\partial(\pi_P^* - \pi_S^*)}{\partial \tau} \geq 0 \tag{11}$$

Proof. Under a pooling equilibrium at a^C, the firm sets $p^* = p_L^* = \delta + \theta_L a^C + \tau$. The profits are given by $\pi_P^* = \delta + \theta_L a^C + \tau - C(a^C)$. Clearly, the derivative with respect to τ is 1.

Under a separating equilibrium, with $a_L^* < a^C$ and $a_H^* < a^C$, the profits are given by $\pi_S^* = q \cdot (\delta + \theta_L a_L - C(a_L^*)) + (1-q) \cdot (\delta + \theta_H(a_H - a_L) + \theta_L a_L - C(a_H^*))$. The derivative of π_S^* with respect to τ is 0. Therefore, $\frac{\partial(\pi_P^* - \pi_S^*)}{\partial \tau} = 1 \geq 0$.

Under a separating equilibrium, with $a_L^* < a^C$ and $a_H^* \geq a^C$, the profits are given by $\pi_S^* = q \cdot (\delta + \theta_L a_L - C(a_L^*)) + (1-q) \cdot (\delta + \theta_H(a_H - a_L) + \tau + \theta_L a_L - C(a_H^*))$. The derivative of π_S^* with respect to τ is $1-q$. Therefore, $\frac{\partial(\pi_P^* - \pi_S^*)}{\partial \tau} = 1 - (1-q) \geq 0$.

Under a separating equilibrium, with $a_L^* \geq a^C$ and $a_H^* \geq a^C$, the profits are given by $\pi_S^* = q \cdot (\delta + \theta_L a_L - C(a_L^*) + \tau) + (1-q) \cdot (\delta + \theta_H(a_H - a_L) + \tau + \theta_L a_L - C(a_H^*))$. The derivative of π_S^* with respect to τ is 1. Therefore, $\frac{\partial(\pi_P^* - \pi_S^*)}{\partial \tau} = 0$.

A.1 Equilibrium Outcomes with a Coarse Certification

A coarse certification influences demand in two ways. First, it creates a discontinuous jump in the marginal valuation of a at the certification requirement, which I note a^C. Second, it lowers the marginal valuation of a for all other values such that θ_i becomes $\hat{\theta}_i$ with $\theta_i \geq \hat{\theta}_i$. These two effects impact the social and monopolist outcomes as follows.

Social Outcome I: $\tilde{\tau}$ is a bias If the parameter $\tilde{\tau}$ is a bias, the perfect information outcome where the marginal valuation of a is set equal to the marginal cost, i.e., $\gamma_i = C'(a_i^*)$, determines the socially optimal level of a for $i = \{L, H\}$. It is never socially optimal to bunch at the certification requirement.

Social Outcome II: $\tilde{\tau}$ is preference If the certification enacts preferences and brings a utility gain of size $\tilde{\tau}$, it can be socially optimal to bunch at the certification requirement.

Lemma 2. *If with certification a_i^* solves $\gamma_i = C'(a_i^*)$ and $a_i^* > a^C$, then the socially optimal level of a for type $i = \{L, H\}$ is: $a_i^{Social} = a_i^*$.*
If a_i^* is the solution of $\gamma_i = C'(a_i^*)$, $a^* \leq a^C$, and $\gamma_H a_i^* - C(a_i^*) \leq \gamma_i a^C + \tilde{\tau}_i - C(a^C)$, then it is optimal to set $a_i^{Social} = a^C$.

Proof. If $\gamma_i = C'(a_i^*)$ and $a_i^* > a^C$, given that the cost function is increasing and convex, social welfare is maximized irrespective of the value of $\tilde{\tau}_i$.
When $\gamma_i = C'(a^*)$ but $a_i^* \leq a^C$, it may be optimal to further increase a_i and to locate a_i at the certification requirement a^C to take advantage of the discrete increase in the willingness to pay: τ_i.

Monopolist Outcome With a coarse certification, where the requirement is set at a^C, and $D_{j=\{L,H\}}$ takes the value 1 if $a_j \geq a^C$ and zero otherwise, the monopolist's problem becomes:

$$\max_{a_H, a_L, p_H, p_L} \pi \cdot (p_L - C(a_L)) + (1 - \pi) \cdot (p_H - C(a_H)) \quad \text{s.t.}$$

$$\text{IRH: } \delta + \hat{\theta}^H a_H + \tau_H D_H - p_H \geq 0$$

$$\text{IRL: } \delta + \hat{\theta}^L a_L + \tau_L D_L - p_L \geq 0$$

$$\text{ICH: } \hat{\theta}^H a_H \tau_H D_H - p_H \geq \hat{\theta}^H a_L + \tau_H D_L - p_L$$

$$\text{ICL: } \theta^L a_L + \tau_L D_L - p_L \geq \hat{\theta}^L a_H + \tau_L D_H - p_H$$

The following algorithm can be used to solve the optimization problem.

1. Solve the relaxed problem using the IR and IC constraints of the monopolist's problem in the presence of a certification. If $a_L^* \geq a^C$ and $a_H^* \geq a^C$, this the optimal solution.

2. If the solution of the relaxed problem is such that: $a_L^* < a^C$ and $a_H^* < a^C$, compare the profits for the following additional two scenarios.

   - Alternative Scenario 1: Set $a_L^* = a^C$ and $a_H^* = a^C$
   - Alternative Scenario 2: Set $a_H^* = a^C$ and solve the relaxed problem for a_L only.

   Between the solution of the relaxed problem and the two alternative scenarios, select the solution with the highest profit.

Proposition 3. *A coarse certification increases the overall provision of quality of a if the following conditions hold:*

1. *without certification:* $a_L^* < a^C$ *and* $a_H^* < a^C$
2. *with certification:* $a_L^* \leq a^C$ *and* $a_H^* = a^C$ *or* $a_L^* = a^C$ *and* $a_H^* \geq a^C$.

Otherwise, the certification decreases the overall provision of quality of a.

Proof. When the certification does not induce the firm to locate a_L and a_H at the certification requirement, the certification decreases the level of a because $\theta_i \geq \hat{\theta}_i$. It is only when the certification induces products to improve their efficiency levels to meet the requirement that the overall provision might increase. This occurs when without certification: $a_L^* < a^C$ and $a_H^* < a^C$, and with certification $a_L^* \leq a^C$ and $a_H^* = a^C$ or $a_L^* = a^C$ and $a_H^* \geq a^C$.

B Additional Summary Statistics and Estimation Details

B.1 Summary Statistics

Table B.1: Market Shares and Model Shares, US Refrigerator Market

	1995	2000	2005	2008
Manufacturer		Market Share		
GE	35%	34%	29%	27%
Electrolux	17%	21%	25%	23%
Whirlpool	27%	24%	25%	33%
Maytag	10%	14%	11%	-
Amana	10%	5%	0%	-
Haier	0%	0%	2%	6%
W.C. Wood	0%	0%	1%	1%
Others	1%	2%	7%	10%
Brand		Model Share		
Kenmore	8%	14%	17%	17%
GE	13%	7%	5%	8%
Kitchen Aid	5%	5%	6%	6%
Amana	8%	4%	3%	3%
Maytag	11%	16%	12%	9%
Whirlpool	7%	5%	10%	10%
Frigidaire	4%	17%	13%	12%
White-Westinghouse	4%	-	-	-
LG	-	-	3%	
Others	50%	36%	30%	27%

Sources: Appliance Magazine; data compiled by the Department of Energy (market share), and California Energy Commission (CEC) Appliance Database (model share). Only full-size refrigerator models on the Californian market for each year are considered. Model shares correspond to the number of models, non-sales weighted, offered by each brand.

C Demand Estimation

The demand model is a latent discrete choice model. For the estimation, the choice probabilities are computed for each household i, zip code r and week t. The probability that household i chooses product j is given by:

$$Q_{irtj} = \sum_{e=\{U,ES,I\}} H_{irt}(e) M_{irtj}(e), \tag{12}$$

Table B.2: Summary Statistics: Retailer's Sample Demand Estimation

	Mean	S.D.
Promotional Price ($)	1311.0	583.7
Manufacturers' Suggested Retail Price ($)	1561.1	703.6
% ES-certified Models	78.9	-
Manufacturers' Reported kWh/y	510.1	74.5
Manufacturers' Reported kWh/y: ES	502.5	68.7
Manufacturers' Reported kWh/y: Non-ES	538.4	88.4
Overall Volume (Cu. Ft.)	22.9	3.1
% More Efficient Minimum Standard	16.5	7.3
Model Share (%) by Door Design		
Top Freezer	25.5	-
Side-by-Side	36.3	-
Bottom-Freezer	38.2	-
Avg # of Refrigerator Models by Zip Code-Trimester	129	45
Total # of Refrigerator Models: Demand Estimation	672	-

Notes: The sample used for the demand estimation consists of all transactions made by homeowners living in single family housing units that bought no more than one refrigerator in the period 2008-2011.

where the term $H_{irt}(e)$ represents the latent probability that consumer i acquires a level of information knowledge such that she is perfectly informed ($e = I$), only informed about ES ($e = ES$), or not informed ($e = U$); and $M_{irtj}(e)$ is the choice probability conditional on the level of knowledge e. The conditional probability $M_{irtj}(e)$ follows a multinomial logit and is defined in Equation 6. The latent probability is also specified as a multinomial logit:

$$H_{irt}(e) = \frac{e^{V_{irt}(e)}}{\sum_k e^{V_{irt}(k)}}, \tag{13}$$

where

$$\begin{aligned} V_{irt}(e = I) =& -K^I - \beta^I X_i + \gamma_1^I MeanElec_{rt} + \gamma_2^I VarElec_{rt} + \gamma_3^I NbModels_{rt} \\ &+ \gamma_4^I VarPrice_{rt}, \\ V_{irt}(e = ES) =& -K^{ES} - \beta^{ES} X_i + \gamma_1^{ES} MeanES_{rt} + \gamma_2^{ES} VarES_{rt} + \gamma_3^{ES} NbModels_{rt} \\ &+ \gamma_4^{ES} VarPrice_{rt}, \\ V_{irt}(e = U) =& 0. \end{aligned} \tag{14}$$

The variables $MeanElec_{rt}$ and $VarElec_{rt}$ are the mean and variance in electricity costs for all products offered in region r at time t, $MeanES_{rt}$ is the proportion of ES-certified products offered, $NbModels_{rt}$ is the number of products in the choice set in a given region, and $VarPrice_{rt}$ is the variance in prices. Note that we have normalized one option of the multinomial logit model such that all coefficients should be interpreted relative to the uninformed type. For instance, a positive coefficient for a variable that enters the type $e = I$ means that such a variable increases the likelihood of being an informed consumer relative to the uninformed type.

The results of the demand estimation are presented below. Panel A presents the estimated coefficients that enter the conditional choice probability $M_{irtj}(e)$. They are the coefficients that define the decision utility of each type. Their interpretation is also presented at the bottom of Panel A. Panel B presents the estimated coefficients that enter the latent probability H_{irt}, where the first set of coefficients are for the informed type ($e = I$), followed by the ES type ($e = ES$). The aggregate share of each type is presented at the bottom of this panel. Note, however, that the shares vary with demographics and across local markets.

Table C.1: Information Acquisition Demand Model

	Income <\$50,000		**Income** ≥\$50,000 & <\$100,000		**Income** ≥\$100,000	
Panel A: Behavioral Parameters Conditional Purchase Decision $M_{irtj}(e)$						
Retail Price (η)	-0.413***	(0.0002)	-0.362***	(0.0001)	-0.317***	(0.0002)
ENERGY STAR τ^{ES}	0.674***	(0.001)	1.528***	(0.002)	1.365***	(0.080)
Rebate (ψ)	0.145***	(0.001)	0.090***	(0.0005)	0.033***	(0.0003)
Elec. Costs (θ)	-4.003***	(0.009)	-3.408***	(0.048)	-4.429***	(0.004)
Interpretation						
Own-Price Elasticity	-5.36		-4.70		-4.12	
Implicit Discount Rate	0.08		0.08		0.03	
WTP ES Label (\$)	163.43		422.22		430.33	
Prob. Taking Rebate	0.35		0.25		0.10	
Panel B: Behavioral Parameters Latent Probabilities $H_{irt}(e)$						
Educ: College ($e = I$)	-0.122***	(0.003)	0.691***	(0.014)	0.303***	(0.012)
Educ: Graduate ($e = I$)	1.717***	(0.031)	2.045***	(0.026)	1.197***	(0.032)
FamSize ($e = I$)	-0.204***	(0.0001)	-0.318***	(0.003)	-0.049***	(0.007)
Age ($e = I$)	0.092***	(0.0002)	0.084***	(0.002)	0.011***	(0.001)
Political: Democrats ($e = I$)	-1.284***	(0.022)	-1.899***	(0.034)	-0.221***	(0.025)
Political: Others ($e = I$)	-1.920***	(0.008)	-1.338***	(0.013)	-0.200	(0.018)
mean-ElecCost ($e = I$)	0.107***	(0.003)	0.075**	(0.001)	0.105***	(0.008)
var-ElecCost ($e = I$)	0.006***	(0.00002)	-0.101***	(0.001)	0.026***	(0.001)
# Models ($e = I$)	0.007***	(0.0001)	0.012***	(0.0001)	0.004 ***	(0.0004)
Variance Price ($e = I$)	-1.003***	(0.004)	-0.729***	(0.012)	-0.390***	(0.004)
Educ: College ($e = ES$)	-0.271***	(0.002)	0.012	(0.007)	0.105***	(0.007)
Educ : Graduate($e = ES$)	-0.453***	(0.014)	0.843***	(0.018)	0.676***	(0.028)
FamSize ($e = ES$)	-0.193***	(0.002)	-0.091***	(0.001)	-0.232***	(0.014)
Age ($e = ES$)	0.063***	(0.0002)	0.045***	(0.001)	0.024***	(0.001)
Political: Democrats ($e = ES$)	-0.255***	(0.006)	-0.421***	(0.015)	-0.045	(0.024)
Political: Others ($e = ES$)	-0.578***	(0.0003)	-0.469***	(0.009)	0.018	(0.025)
Proportion-Estar ($e = ES$)	2.837***	(0.002)	0.975***	(0.001)	2.324***	(0.114)
# Models ($e = ES$)	-0.006***	(0.0002)	-0.001***	(0.0000)	-0.003***	(0.001)
Variance Price ($e = ES$)	0.316***	(0.004)	0.211***	(0.004)	0.109***	(0.006)
Interpretation						
$H(e = I)$	0.34		0.50		0.56	
$H(e = ES)$	0.21		0.10		0.17	
$H(e = U)$	0.45		0.41		0.27	
# Obs.	46,097		45,487		45,249	
LLE	188,088		194,394		195,969	

Notes: Asymptotic robust standard errors in parentheses: * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$). Prices, rebates, and electricity costs measured in hundreds of dollars. Average price of \$1,300 used to compute own-price elasticity. Refrigerator lifetime of 18 years used to compute implicit discount rate.

C.1 Construction of the Representative Choice Set

To construct the representative choice set for the retail costs estimation and policy simulation, I sample refrigerator models used in the demand estimation. I sample refrigerator models to match the distribution of important attributes of the choice set offered on the US market in 2011. To illustrate, suppose that 5% of the full-size refrigerators available on the US market in 2011 were GE top-freezer refrigerators between 16 cu.ft. and 21 cu.ft. and certified ES. I sampled 3 refrigerator models in my sample (5% X 68 $\approx$ 3) that fit this description. I use the estimated product fixed effects to determine the location of each product in the quality dimension, which I hold fixed throughout the estimation and policy simulation.

Table C.2: Summary Statistics Representative Choice Sets:
FTC versus Supply Estimation/Policy Simulation

	Observed FTC 2011		Constructed Retailer's Sample	
	Mean	S.D.	Mean	S.D.
Model Share (%) by Brand				
A	16.7		10.3	
B	19.2		22.1	
C	15.5		19.1	
D	8.5		10.3	
E	20.6		11.8	
F	19.5		26.5	
Model Share (%) by Door Design				
Top Freezer	32.5		22.1	
Side-by-Side	37.9		41.2	
Bottom-Freezer	29.7		36.8	
Overall Volume (Cu. Ft.)	22.0	3.4	23.6	2.8
Manufacturers' Reported kWh/y	507.5	91.5	514.0	74.2
% Certified ES	58.4		67.7	
% More Efficient Minimum Standard	17.2	7.0	18.3	7.4
# Models	1,828		68	

Notes: The FTC provides data for all refrigerator models offered on the market for the year 2011. The first two columns report the mean standard deviation for various attributes "observed" in the FTC data. The ES certification status of each model offered was added using data from the EPA. The "constructed" choice set consists of a random sample of refrigerator models draws from the set of models offered by the retailer and used in the demand estimation. All values reported are not sales-weighted. The constructed choice set is used for both the estimation of the unit retail costs and the policy simulations.

C.2 Emission Factors

Table C.3: Emission Factors and Externality Costs

Non-baseload Output Emission Rates (US Average)			
Pollutant	Estimate		Source
CO_2	1583 lb/MWh		
CH_4^a	35.8 lb/GWh		
N_2O^a	19.9 lb/GWh		EPA, eGRID2007
SO_2	6.13 lb/MWh		
NOx	2.21 lb/MWh		
Damage Cost (2008 $)			
Pollutant	Low Estimate	High Estimate	Source
CO_2	$21.8/t	$67.1/t	Greenstone et al. (2013)
SO_2	$2060/t	$6700/t	low: Muller and Mendelsohn (2012), high: EPA[b]
NOx	$380/t	$4591/t	low: Muller and Mendelsohn (2012), high: DOE[c]

Notes: (a) Externality costs associated with CH_4 and N_2O are assumed to be the same as for CO_2. CH_4 and N_2O are converted in CO_2 equivalent using estimates of global warming potential (GWP). The GWP used for $CH4$ is 25, and the GWP used for N_2O is 298. Source: IPCC Fourth Assessment Report: Climate Change 2007. (b) Estimates used in the illustrative analysis of the 2012 regulatory impact analysis for the proposed standards for electric utility generating units. (c) Higher value of the estimate used in the federal rule for new minimum energy efficiency standards for refrigerators (1904-AB79).

D Policy Analysis: Simulation Details and Additional Results

D.1 Welfare Measure

In Houde (2018), I show that under the assumption that the decision utility of the informed consumer type coincides with experienced utility, the demand model provides a measure of compensating variation (CV) for a policy change $\mathcal{P} \rightarrow \tilde{\mathcal{P}}$ based on the concept of experienced utility:

$$
\begin{aligned}
CV_i = \frac{1}{\eta}\Bigg\{ & \tilde{H}_i^I \cdot ln \sum_j^J exp(\tilde{U}_{ij}^I) - h_i^I \cdot ln \sum_j^J exp(U_{ij}^I) \\
& + \tilde{H}_i^{ES} \cdot \left[ln \sum_j^J exp(\tilde{U}_{ij}^{ES}) + \sum_j^J \tilde{M}_i^{ES} \cdot (\tilde{U}_{ij}^I - \tilde{U}_{ij}^{ES}) \right] \\
& - H_i^{ES} \cdot \left[ln \sum_j^J exp(U_{ij}^{ES}) + \sum_j^J M_i^{ES} \cdot (U_{ij}^I - U_{ij}^{ES}) \right] \\
& + \tilde{H}_i^U \cdot \left[ln \sum_j^J exp(\tilde{U}_{ij}^U) + \sum_j^J \tilde{M}_i^U \cdot (\tilde{U}_{ij}^I - \tilde{U}_{ij}^U) \right] \\
& - H_i^U \cdot \left[ln \sum_j^J exp(U_{ij}^U) + \sum_j^J M_i^U \cdot (U_{ij}^I - U_{ij}^U) \right] \Bigg\}.
\end{aligned} \tag{15}
$$

Applying the above formula to each income group, we can obtain an income-specific measure of CV, which considers that some consumers may make a purchase decision without complete information and may be subject to behavioral biases. This welfare measure departs from the standard expression for logit-based discrete choice models (Small and Rosen 1981) in two ways. First, it has the correction term $\sum_j^J M_i^{ES,U}(U_{ij}^I - U_{ij}^{ES,U})$, which captures the expected difference between experienced and decision utility for relying on the ES certification ($e = ES$), or for being uninformed ($e = U$), instead of being fully informed ($e = I$). This term captures the magnitude of the misperceptions due to imperfect information and behavioral biases. For the case where $e = ES$, the size of the misperception is partly induced by the label effect (parameter τ), which captures the large willingness to pay for ES-certified products that goes well beyond average energy savings. In Equation 15, the parameter τ enters U_{ij}^{ES}, but does not enter U_{ij}^I, which means that the label effect acts as a bias and does not influence experienced utility. Therefore, in a scenario without

ES, consumer welfare will not mechanically decrease because the label is not present. My welfare measure thus provides a conservative estimate of the benefits consumers derive from ES.

The second difference between the expression in Equation 15 and the standard measure of welfare is that the overall CV is a weighted sum of the CV experienced by different latent consumer types, where the weights are the probabilities H^e. The expression thus allows me to decompose the overall change in consumer welfare and report the incidence of a policy change on different types of consumers, although they are not readily observed. For instance, in the present application, I can report how the ES program impacts uninformed versus informed consumers.

D.2 Simulation Details

I simulate two scenarios to quantify the welfare effects of the ES certification. In one scenario, I simulate an equilibrium with the ES certification. For the main policy simulation, the certification requirement is set relative to the federal minimum energy efficiency standard that was in effect in 2011. The minimum standard represents a constraint on firms' strategies with respect to the kWh/y offered for a particular model. Each refrigerator model in the choice set has a minimum standard defined with respect to the size of the refrigerator and freezer location. In this scenario, a model is certified if it meets or exceeds the certification requirement. That is, the certification process is assumed to be costless. In the second scenario, I simulate an equilibrium without the ES certification.

I simulate the demand model using a subsample of households used for the demand estimation. Given that the size of the subsample has a notable impact on the computation time, I select a subsample of 3,500 households, which I found sufficient to obtain representative distributions of demographics and local electricity prices.

I must also fix the quality in the non-energy dimension to simulate the demand model. I use the estimated product fixed effects for that purpose and construct quality terms in the non-energy dimension as follows. Given that the demand model was estimated separately for the three income groups, there is a set of product fixed effects specific to each income group. Therefore, the perceived quality in the non-energy dimension differs across income groups.

For each scenario, I solve for the Nash equilibrium where firms decide the price and kWh/y of each model they offer. I use the Gauss-Seidel best-response iteration algorithm, which consists of solving for the optimal combination of prices and kWh/y for one firm, holding other firms' strategies fixed, and iterating over each firm until the strategies converge to a fixed point. I use the change in profits from one iteration to the other as the convergence criterion. For each iteration, I also randomize the order in which each firm determines its strategy. If the algorithm takes five iterations, the order in which each firm determines its optimal strategy varies across these five iterations. For instance, sometimes Brand A may play first; sometimes it may play last.

In this game, the non-existence of an equilibrium and multiple equilibria are possible. Non-existence is rarely an issue in practice. However, multiple equilibria are present, especially in the scenario with ES. The fact that the certification creates a discontinuity in the valuation of energy efficiency induces oscillation in the optimal strategies across iterations. The algorithm takes a much higher number of iterations to converge with ES relative to without ES. For instance, for a scenario without ES, the Gauss-Siedl algorithm usually converges after 5 iterations. The scenario with ES converges after 40 iterations, on average. There are multiple equilibria in the scenario without certification, but they are very close. The differences are more important with ES, but not economically or statistically significant. Figure D.1 first makes this point and shows the trajectory of each product in the energy efficiency-price space for 6 different simulations.[28] The starting point of each arrow corresponds to the energy efficiency level and price of one particular product under a scenario with ES. The endpoint corresponds to the location of this same product when ES is removed. We can readily see that the trajectories tend to converge to the same point, which shows that the equilibrium without ES, although not unique, is very stable. The starting points, i.e., equilibrium with ES, are more scattered, but they are consistently in the same region of the product space. They are thus at the source of uncertainty. Table D.1 quantifies this uncertainty. I simulate the model 25 times, with and without ES, and report the mean and standard errors across simulations for each metric.

[28] I keep the number of simulations relatively low to have a tractable figure.

Compared to the mean estimates, the standard errors are small, less than 10% in most cases. The simulation results, and more importantly, the conclusions I draw from them, are thus robust to the existence of multiple equilibria. Therefore, I do not implement an equilibrium selection procedure for the different counterfactual scenarios presented in the main text. I, however, acknowledge some uncertainty in the magnitude of the estimates that stems from having multiple equilibria.

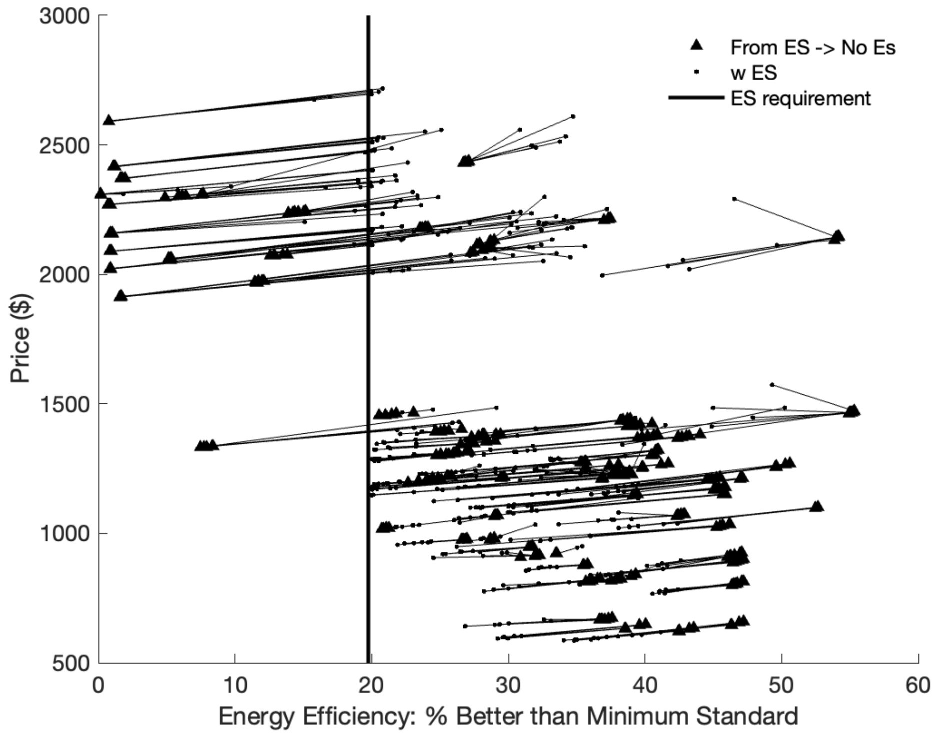


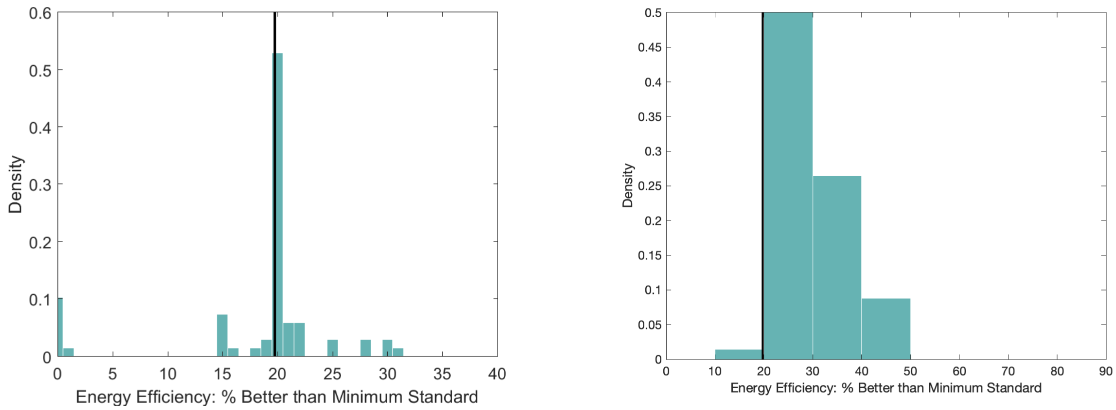
Figure D.1: Multiple Equilibria: Energy Efficiency and Price Offered
Notes: Each arrow represents the movement in the energy efficiency-price space with ES (starting point) and without ES (end point). The figure presents the trajectories for 6 different simulations to show that there are multiple equilibria close to each other.

Table D.1: Uncertainty Multiple Equilibria: The Effects of Removing the ES Certification

	Income <\$50,000	Income ≥\$50,000 & <\$100,000	Income ≥\$100,000	All
Δ CV	19.1	9.0	17.0	14.0
	(0.9)	(1.3)	(1.5)	(1.2)
Δ CV, $e = I$	21.3	16.9	22.7	19.7
	(0.7)	(1.2)	(1.3)	(1.0)
Δ CV, $e = U$	2.0	-8.3	-31.0	-11.6
	(0.7)	(1.1)	(1.6)	(1.1)
Δ Externalities-Low	-8.1	-1.4	-2.3	-3.5
	(0.3)	(0.3)	(0.2)	(0.2)
Δ Externalities-High	-26.5	-4.7	-7.5	-11.6
	(0.9)	(0.8)	(0.7)	(0.8)
Δ Profits				-2.9
				(0.4)
Δ Social Welfare-Low				14.5
				(1.4)
Δ Social Welfare-High				22.6
				(1.9)

Notes: The table reports the difference between a market without and with ES. The counterfactual scenario is the market without ES. A negative sign implies that removing ES leads to a decrease in a particular metric relative to a market with ES. The model is simulated 25 times to investigate multiple equilibria. The standard errors (in parentheses) represent the variation across equilibria. Compared to the mean estimates, the standard errors are small—less than 10% in most cases. The term CV refers to the compensating variation and quantifies the change in consumer welfare. The compensating variation for the informed type ($e = I$) and uninformed type ($e = U$) are also presented. As before, the externalities are computed for two scenarios: a low and high range of estimates for the marginal damage costs.

D.3 Observed versus Simulated Distributions



(a) Observed (2011), with ES

(b) Predicted, with ES

Figure D.2: Observed versus Predicted Distributions of Energy Efficiency
Notes: Panel (a) is the distribution of products in the energy efficiency space observed for the representative choice set of 68 products. Panels (b) presents the simulated distribution, where the requirement is set at 20% more stringent than the federal minimum energy efficiency standard.

D.4 Heterogeneity in the Cost Function

I consider heterogeneity in the cost function by interacting the measure for energy efficiency in the pair-matching estimator with a dummy for specific attributes. I investigate two dimensions of heterogeneity in separate regressions. The model I use takes the form:

$$ln(price_{j,r}^{manuf}) = \alpha + \gamma_{j,j'} + Y_j \times Brand_j + \phi Efficiency_j + \psi Efficiency_j \times D_{attribute} + \epsilon_{j,r}. \quad (16)$$

In one specification, the dummy distinguishes top-freezer refrigerators from other designs (i.e., side-by-side and bottom-freezer refrigerators). In a second specification, the dummy distinguishes larger refrigerators ($\geq$ 20 Cu. Ft.) from smaller ones ($<$ 20 Cu.Ft). The results in Table D.2 suggest that side-by-side/bottom-freezer refrigerators or larger refrigerators have a lower marginal cost to improve energy efficiency. These estimates are, however, imprecise and not statistically significant at the conventional level.

If I consider heterogeneity in the cost function, the welfare results (Table D.3) are very similar. For instance, if we take the estimates from the pair-matching estimator where we distinguish top-freezers refrigerators from other designs (second specification with year-brand FE) and use this cost function for the simulation, we found results that are qualitatively similar to the estimates presented in the main text. The model with the heterogeneous cost function tends to predict less bunching at the ES requirement (see Figure D.3). It is expected, given that for some models, the marginal cost of energy efficiency is now lowered. In the absence of ES, firms offer less inefficient models as well as a larger share of highly efficient models.

Table D.2: Paired Refrigerator Models: Estimation Results with Heterogeneity

Estimation Marginal Cost Providing EE (ϕ) and interaction term (ψ):

$$ln(price^{manuf}_{j,r}) = \alpha + \gamma_{j,j'} + Y_j \times Brand_j + \phi Efficiency_j + \psi Efficiency_j \times D_{attribute} + \epsilon_{j,r},$$

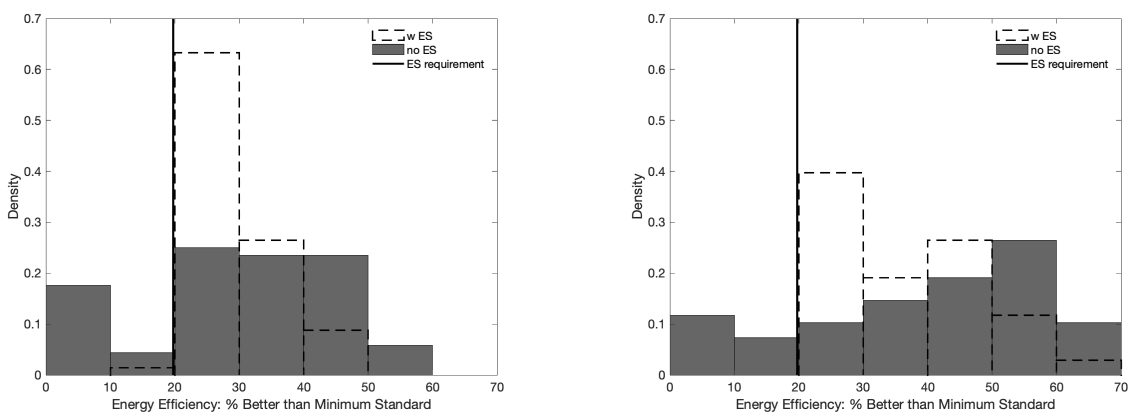
Panel A: Heterogeneity w.r.t to Freezer Location		
	ϕ: Eff. of Top-Freezer	ψ: Eff. $\times$ Other Designs
Pair FE only	184.1	-11.2
	(60.2)	(165.1)
Pair FE & Year-Brand FE	203.9	-103.9
	(72.7)	(203.7)
Hedonic Regression	230.8	-164.0
	(101.2)	(362.7)
Panel B: Heterogeneity w.r.t to Overall Size		
Attribute	ϕ: Eff. of Smaller Size (<20 Cu.Ft.)	ψ: Eff. $\times$ Larger Size ($\geq$20 Cu.Ft.)
Pair FE only	189.5	-64.0
	(59.2)	(180.4)
Pair FE & Year-Brand FE	202.2	-103.7
	(71.9)	(210.5)
Hedonic Regression	222.5	-166.0
	(100.6)	(312.6)

Notes: This table reports heterogeneity in the marginal cost of providing energy efficiency along two dimensions. Panel A considers heterogeneity with respect to the freezer location. The main estimate, ϕ, corresponds to the marginal cost of energy efficiency for top-freezer refrigerators. The second column is an interaction term, ψ, for side-by-side and bottom-freezer refrigerators. A negative term for ψ means that the marginal cost is lower relative to the main effect. Panel B considers heterogeneity with respect to overall size, where a dummy for refrigerators bigger than 20 Cu. Ft. is interacted with the measure of energy efficiency. A negative sign on the interaction term means that larger refrigerators have a lower marginal cost of improving energy efficiency. The sample used for these regressions is the same as the one used in the main text for the identical pairs of refrigerator models. Standard errors are in parentheses.

Table D.3: The Effects of Removing the ES Certification with Heterogeneous Marginal Cost

	Income <\$50,000	Income ≥\$50,000 & <\$100,000	Income ≥\$100,000	All
Δ CV	15.2	5.6	22.4	12.8
Δ CV, e=I	25.2	18.1	34.3	24.5
Δ CV, e=U	4.4	-7.6	-22.3	-8.2
Δ Externalities-Low	-7.3	-3.9	-8.0	-5.9
Δ Externalities-High	-24.2	-12.7	-26.2	-19.6
Δ Profits				-2.9
Δ Social Welfare-Low				15.9
Δ Social Welfare-High				29.5

Notes: The table reports the difference between a market without and with ES. The counterfactual scenario is the market without ES. A negative sign implies that removing ES leads to a decrease in a particular metric relative to a market with ES. All figures are in dollars per consumer. The term CV refers to the compensating variation and quantifies the total change in consumer welfare. The compensating variation for the informed type ($e = I$) and uninformed type ($e = U$) are also presented. The negative externalities associated with electricity consumption are evaluated for two scenarios: 'Low' refers to the lower range of the damage estimates, and 'High' refers to the higher range of the estimates. For both scenarios, the dollar value of the negative externalities decreases without certification, while the total welfare increases.



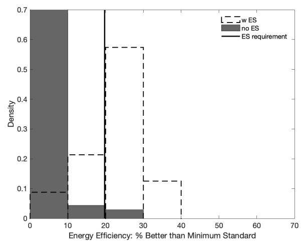
(a) Homogeneous Cost Function

(b) Heterogeneous Cost Function

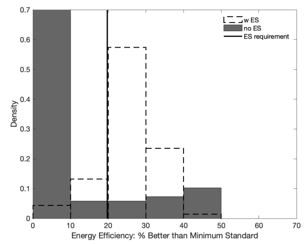
Figure D.3: Distributions of Energy Efficiency for Different Cost Functions

Notes: Each panel presents the distributions of energy efficiency with and without ES. The distributions simulated with the heterogeneous cost function predict more differentiation in the absence of ES.

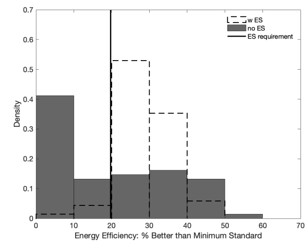
D.5 Share of Informed/Uninformed Consumers



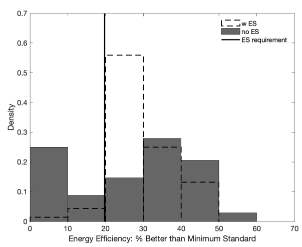
(a) 9%/67% → 12%/88%



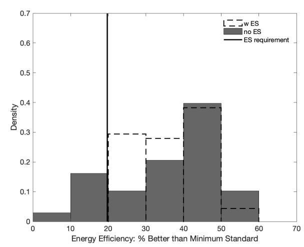
(b) 21%/58% → 27%/73%



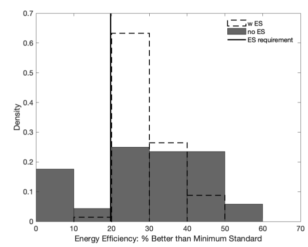
(c) 31%/51%→ 38%/62%



(d) 42%/42% → 50%/50%



(e) 67%/24%



(f) Base-case

Figure D.4: Distributions of Energy Efficiency With and Without ES for Different Shares of Informed/Uninformed Consumers

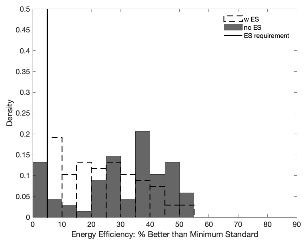
Notes: Each panel plots the distributions of energy efficiency with (dotted) and without (dark) the ES certification for different shares of informed/uninformed consumers. The percentages represent the share with and without ES, respectively. The last panel is the base-case scenario where the share of informed/uninformed consumers with ES is 35%/36% and 56%/44% without ES. The share of ES type can be inferred from the first two percentages. To illustrate, in the base-case scenario it is: 100%-35%-36%=29%.

Table D.4: The Effects of Removing the ES Certification as a Function of the Shares of Informed/Uninformed Consumers

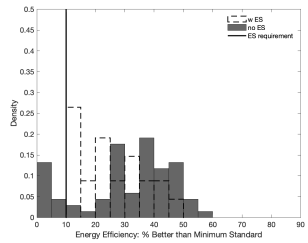
	Sensitivity Tests					Base-case
	I	II	III	IV	V	
Panel A: Average Share of Informed/Uninformed Consumers						
Share of Informed with ES: $H(e = I)$	0.09	0.21	0.31	0.42	0.67	0.35
Share of Uninformed with ES: $H(e = U)$	0.67	0.58	0.51	0.42	0.24	0.36
Share of Informed no ES: $H(e = I)$	0.12	0.27	0.38	0.50	0.73	0.56
Share of Uninformed no ES: $H(e = U)$	0.88	0.73	0.62	0.50	0.27	0.44
Panel B: Welfare Metrics						
Δ CV	-75.1	-58.3	-36.5	-1.3	-4.3	12.4
0.0 Δ CV, Income <\$50,000	-76.8	-57.6	-28.0	6.8	0.3	18.8
"Δ CV, Income ≥\$50,000 & <\$100,000"	-55.6	-43.7	-25.4	2.3	-2.6	7.0
Δ CV, Income ≥\$100,000	-105.8	-83.3	-63.7	-15.8	-11.8	15.0
Δ Externalities-Low	31.1	24.0	16.3	4.4	1.6	-3.4
Δ Externalities-High	102.3	79.1	53.6	14.4	5.2	-11.3
Δ Profits	-8.5	-7.0	-8.3	-2.5	-1.2	-5.2
Δ Social Welfare-Low	-114.7	-89.3	-61.0	-8.2	-7.0	10.7
Δ Social Welfare-High	-185.9	-144.3	-98.4	-18.3	-10.6	18.6

Notes: The table reports the difference between a market without and with ES. The counterfactual scenario is the market without ES. In the five sensitivity scenarios, the shares of informed and uninformed consumers do not vary with demographics or choice set variables, unlike in the base case. The probability $H(e = I)$ corresponds to the share of informed consumers with or without ES. The probability $H(e = U)$ is the share of uninformed consumers. When ES is in effect, the share of consumers that rely on ES can be inferred by summing $H(e = I)$ and $H(e = U)$. The last column shows the results for the base-case scenario reported in the main text. Note that the shares of informed and uninformed consumers vary across groups in the base-case scenario, but I only report the average across the three income groups. In the sensitivity scenarios, the share of informed/uninformed consumers is equal across the three income groups.

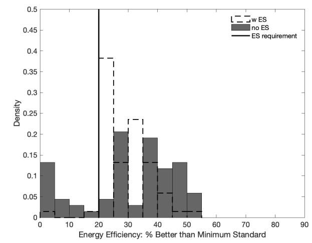
D.6 Optimal Stringency Requirement



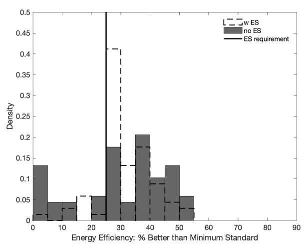
(a) 5%



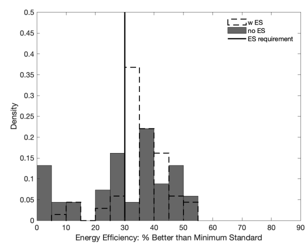
(b) 10%



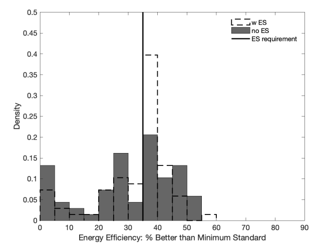
(c) 20%



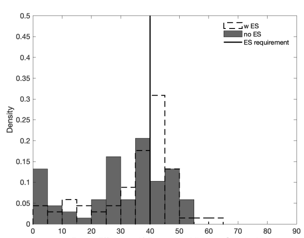
(d) 25%



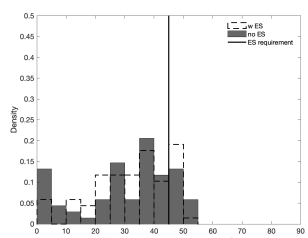
(e) 30%



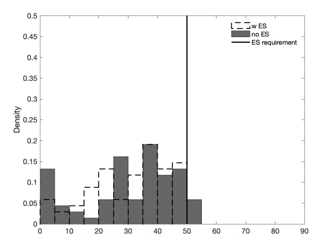
(f) 35%



(g) 40%



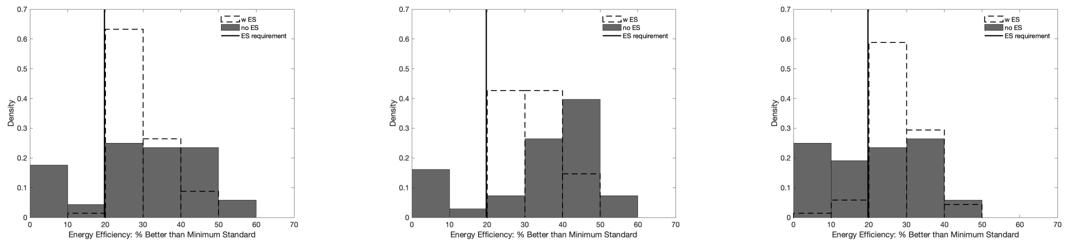
(h) 45%



(i) 50%

Figure D.5: Distributions of Energy Efficiency and Certification Stringency
Notes: Each panel plots the distributions of energy efficiency with (dotted) and without (dark) the ES certification for a given stringency of the certification requirement. For stringency requirements ranging from 10% to 30%, the mass of the distribution without certification is more important at high efficiency levels.

D.7 Interaction with Electricity Prices



(a) Base-case (b) Average Prices BB (c) SMC Prices BB

Figure D.6: Distributions of Energy Efficiency and Electricity Prices
Notes: Each panel plots the distributions of energy efficiency with (dotted) and without (dark) the ES certification for a given scenario regarding the electricity prices households pay. Panel (a) corresponds to the base case where households pay electricity prices that correspond to county averages. Panel (b) corresponds to a scenario where households pay electricity prices that correspond to county averages from the sample period of Borenstein and Bushnell (2022): 2014-2016. Panel (c) corresponds to a scenario where households pay electricity prices set at the social marginal costs calculated by Borenstein and Bushnell (2022).

Table D.5: Change in Consumer Surplus Across US Regions with Low, Medium, or High Electricity Prices

	<\$50,000	≥\$50,000 & <\$100,000	≥\$100,000	All
Δ CV	18.8	7.0	15.0	12.4
Δ CV, Low Elec. Price: < 0.11 \$/kWh	7.6	4.8	8.8	6.7
Δ CV, Medium Elec. Price: 0.11 to 0.16 \$/kWh	19.4	6.2	12.8	11.7
Δ CV, High Elec. Price: > 0.16 \$/kWh	65.1	21.0	54.4	42.4

Notes: The table reports the difference in consumer surplus between a market without and with ES. The first row is an average across all regions of the United States. The other rows present the change in consumer surplus for different regions based on average county electricity prices. Regions with high electricity prices benefit the most from removing the ES certification, because of the increase in high efficiency models offered in this scenario.

Table D.6: The Effects of Removing the ES Certification as a Function of Electricity Prices

	Electricity Price (Faced by All Households)		
	Avg. Price 2011	Avg. Price BB	SMC BB
Δ CV	12.4	17.0	1.1
Δ CV, Income <$50,000	18.8	27.0	3.5
Δ CV, Income ≥$50,000 & <$100,000	7.0	7.3	1.3
Δ CV, Income ≥$100,000	15.0	22.7	-2.0
Δ Externalities-Low	-3.4	-7.3	7.4
Δ Externalities-High	-11.3	-24.2	24.3
Δ Profits	-5.2	-3.2	-3.4
Δ Social Welfare-Low	10.7	21.2	-9.7
Δ Social Welfare-High	18.6	38.0	-26.6

Notes: The table reports the difference between a market without and with ES. The counterfactual scenario is the market without ES. A negative sign implies that removing ES leads to a decrease in a particular metric relative to a market with ES. The first column represents a scenario where the electricity prices paid by households are the 2011 county average electricity prices. The second column uses the county average electricity prices computed by Borenstein and Bushnell (2022) for the period 2014-2016. The third column households pay electricity prices set at the social marginal costs as determined by Borenstein and Bushnell (2022).