Title
Do Energy Efficiency Standards Improve Quality? Evidence from a Revealed Preference Approach:

Authors
Houde, Sebastien
Spurlock, C. Anna

Publication Date
2015

LBNL-182701

ERNEST ORLANDO LAWRENCE BERKELEY NATIONAL LABORATORY

Do Energy Efficiency Standards Improve Quality? Evidence from a Revealed Preference Approach

Sébastien Houde*, C. Anna Spurlock†

*Department of Agricultural and Resource Economics
University of Maryland

† Energy Analysis and Environmental Impacts Division
Lawrence Berkeley National Laboratory

June 2015

This work was supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Do Energy Efficiency Standards Improve Quality? Evidence from a Revealed Preference Approach

Sébastien Houde,* C. Anna Spurlock†

June 1, 2015

Abstract

Minimum energy efficiency standards have occupied a central role in U.S. energy policy for more than three decades, but little is known about their welfare effects. In this paper, we employ a revealed preference approach to quantify the impact of past revisions in energy efficiency standards on product quality. The micro-foundation of our approach is a discrete choice model that allows us to compute a price-adjusted index of vertical quality. Focusing on the appliance market, we show that several standard revisions during the period 2001-2011 have led to an increase in quality. We also show that these standards have had a modest effect on prices, and in some cases they even led to decreases in prices. For revision events where overall quality increases and prices decrease, the consumer welfare effect of tightening the standards is unambiguously positive. Finally, we show that after controlling for the effect of improvement in energy efficiency, standards have induced an expansion of quality in the non-energy dimension. We discuss how imperfect competition can rationalize these results.

Keywords: Minimum Standards, ENERGY STAR, Energy Efficiency, Product Differentiation.

*Department of Agricultural and Resource Economics, University of Maryland, shoude@umd.edu.

†Lawrence Berkeley National Lab, caspurlock@lbl.gov. This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Building Technology, State, and Community Programs, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

1. Introduction

Minimum energy efficiency standards have occupied a central role in U.S. energy policy for more than three decades. They were first introduced as a response to the 1970s energy crisis. At that time, the economic rationale (Hausman and Joskow 1982) was that energy prices paid by consumers were well below marginal costs as well as the apparent fact that consumers overdiscount future energy costs (Hausman 1979). More than 30 years later, minimum energy efficiency standards are more popular among policy-makers than ever. The Obama administration has relied heavily on standard-based policies in the president's Climate Action Plan. As a result, the depth and breadth of the coverage of these policies in the markets of energy intensive durables have substantially increased and will continue to do so in the next few years.[1]

Despite their prominence, little is known about the welfare effects of current energy efficiency standards, especially outside the passenger vehicle market, and there rationales are still debated (Parry, Evans, and Oates 2014). Standard rulemaking analyses provide ex ante estimates of the costs and benefits of new and revised standards to demonstrate that they are "economically justified," but there have been few ex post analyses investigating how existing standards for energy intensive durables have actually performed.[2]

The goal of this paper is to show how standards have impacted overall quality in the appliance market. We employ a revealed preference approach that allows us to compute a price-adjusted quality index. This quality index, together with information on prices, is sufficient to determine the consumer welfare effect of past standards. Our empirical analysis focuses on six appliance

[1]New minimum standards are established and revisions of existing standards are set into law a few years before they become effective. The Department of Energy (DOE) under the current Obama administration has enacted numerous standards that will become effective after 2016.

[2]Dale, Antinori, McNeil, McMahon, and Fujita (2009) retrospectively looked at price trends in the appliance market and found that following the revisions in minimum standards, realized prices were much lower than the prices anticipated in the rulemaking analyses. Spurlock (2014) showed that appliance standards led to a decrease in appliance prices at the time the revisions became effective. Allcott and Taubinsky (2013) used a choice experiment to investigate whether a ban on incandescent lightbulbs, similar to one recently implemented in California, was justified, and concluded that it was most likely not. Brucal and Roberts (2015) is a study complementary to ours that uses the same data. They also look at the evolution of price and quality for one of the six appliance categories that we consider: the U.S. clothes washer market. They develop a quality index using a different methodology and find results similar to ours.

categories[3] during the the 2001-2011 period, which were subject to frequent revisions in two types of energy efficiency standards: minimum and ENERGY STAR (ES) standards. Minimum standards mandate maximum energy use for all products manufactured in a given year and the ES program is a voluntary certification that complements minimum standards by identifying the most energy efficient products in the marketplace. By exploiting variation in both minimum and ES standards, we are able to show how products located at different ends of the energy efficiency spectrum respond to standards.

We find that during the period 2001-2011, the price-adjusted quality indexes of six different appliance categories either increased or remained constant. We also observe several discrete increases in quality that coincide with the effective date of new or revised standards—-newly introduced appliance models that met the more stringent standards were then of higher quality, on average, relative to existing models on the market. We propose four estimators to quantify the change in quality and price of appliance models that were marginal to the standards. All four estimators exploit the fact that manufacturers are strategic in the timing of their product line decisions and compare appliance models that exited the market at the time of a standard revision to models that entered the market during this same period. We argue and show evidence that these models identify the marginal effect of a standard revision. We present two estimators that control for unobserved temporal shocks by using variation in the timing of standard revisions across appliance categories. Finally, we also implement a matching estimator that distinguishes the change in quality due to incremental improvement of existing product lines versus the introduction of new product lines subject to major redesign. The results suggest that most standard revisions in our sample led to an increase in overall quality. We, however, find a modest impact on price; some revision events led to a decrease in prices, but most price changes are not statistically significant. We also find that the change in quality is due to both the introduction of new product lines and improvement within product lines.

The micro-foundation of our quality index is a discrete choice model for differentiated products. Our econometric procedure is inspired by hedonic models used in industrial organization (Ackerberg and Rysman 2005; Bajari and Benkard 2005; Berry and Pakes 2007; Song 2007). Our main estimator consists of a nested logit that explicitly accounts for the effect of unobserved horizontal product differentiation with the addition of a function in the alternative-specific utility that counts the

[3]In the main analysis, we present results for clothes washers, dishwashers, refrigerators (both full sized and compact), freezers, and room air conditioners without reverse cycle. In the online Appendix, we present additional results for room air conditioners with reverse cycle.

number of products and provides a measure of distance between products, as proposed by Ackerberg and Rysman (2005). We estimate the model using monthly national market shares disaggregated at the product level. These data are not suited to identify rich heterogeneity patterns in consumer preferences as is typically done in demand estimation for differentiated products (Berry, Levinsohn, and Pakes 1995, 2004). We focus on identifying a measure of vertical quality that captures how each product is valued in equilibrium on the market. Our approach and data are similar in nature to what has been used in the trade literature (Hallak and Schott 2011; Feenstra and Romalis 2014). One important difference from this literature is that our fine level disaggregation at the product level allows us to side-step aggregation issues in the estimation and construction of indexes.

We show that our results are robust to our modeling assumptions through extensive sensitive tests. For instance, instead of estimating the coefficient on price, we calibrate it and perform sensitivity tests that cover a wide range of economically plausible values for this parameter. For any value of this parameter, the results are qualitatively the same—several standard revisions are associated with a large increase in price-adjusted quality.

The micro-foundation of our approach has several advantages. The quality index has a precise economic interpretation, which is simply the so-called mean utility term in the model of Berry, Levinsohn, and Pakes (1995, 2004). Our quality index can then be decomposed to identify which product characteristics contribute to the overall change in quality. In the present application, we are particularly interested in determining whether changes in overall quality are driven primarily by energy use and how standard revisions have affected quality in the non-energy dimension. Previous work has suggested that prices of appliance models that remained in the market across standard changes dropped, while average prices remained unchanged (Spurlock 2014). The concern is that manufacturers might have been trading off quality of the products against improvements in energy use, in order to maintain their price points. If this were the case, then we might expect to see quality in the non-energy dimension to diminish. The opposite is also possible. Several theoretical papers (e.g., Ronnen 1991; Crampes and Hollander 1995) investigating quality standards in imperfectly competitive markets have shown that standards could simultaneously induce lower prices and improve quality beyond the required minimum. The intuition behind these results is that standards reduce product differentiation in the regulated dimension of the product space, which increases competition among products and incentivizes firms to further differentiate by expanding quality. Despite the extensive theoretical literature on quality standards and market structure, very little empirical work has been done to test this hypothesis.[4] We then propose two approaches to test how energy efficiency standards impact quality in the non-energy dimension.

We first propose a lower bound on an energy efficiency-price-adjusted quality index. This index is a conservative estimate of quality in the non-energy dimension that we compute by making a generous adjustment to how consumers value future energy costs. Using this index, we find evidence that quality in the non-energy dimension also increased following the revision of standards. In a second approach, we directly estimate how consumers value future energy costs. We focus on one appliance category (clothes washers) as a case study and use rich and detailed data on product characteristics collected from product users' manuals. Given the extensive list of product characteristics that we observe, not all which may be valued by consumers, our challenge is to deal with a sparse demand model. Similarly to Gillen, Shum, and Moon (2014), we propose a LASSO regression to select and measure how product characteristics correlated with the price-adjusted quality index. Using this procedure, we find that the coefficient on energy use is negative and statistically significant, but the estimate suggests a considerable undervaluation of future energy costs by consumers. We also find that clothes washers' overall capacity is highly valued by consumers. We show that the number of features offered by clothes washer manufacturers have drastically increased during the sample period and manufacturers have relied more on trademarked technologies over time. More importantly, we observe discrete changes in the rate of increase in the number of features and trademarks that coincide with the effective date of some standard revisions. The LASSO regression selects several of these characteristics and from the characteristics selected about half are statistically significant. This suggests that from the large and increasing number of features offered during this period, several were valued by consumers. However, once we account for improvement in energy use and overall capacity through time, the remaining component of quality of clothes washers is no more increasing, but decreasing over the sample period. In sum, lower energy use and larger size clothes washers appear to be the main drivers in the improvement in overall quality, not the featurization and trademarking of the products.

Our conclusions have important implications for the design of energy efficiency standards. The current paradigm is that standards lead to overall improvement in energy efficiency, but induce

[4]Examples we could find were on quality regulation or licensing for certain service industries (e.g., child care centers: Gormley 1991; Chipty and Witte 1997; professional services: Carroll and Gaston 1981, 1983; and drug introductions following FDA regulation: Wiggins 1981).

higher prices, reduce product diversity (Hausman and Joskow 1982), and may lower quality in other dimensions.[5] In several instances, we find exactly the opposite.

The remainder of this paper is organized as follows. In the next section, we give an overview of the policy background related to appliance standards. In Section 3, we discuss the related literature and predictions from theoretical papers. In Section 4, we present our approach used to construct a quality index. In Section 5, we discuss our data and estimation strategy. In Section 6, we present the results for the overall evolution of quality in the appliance market and regression results quantifying the change in quality and price. In Section 7, we present two approaches to investigating the evolution of quality in the non-energy dimension. We conclude in Section 8.

2. Policy Background

U.S. federal energy efficiency regulation of energy consuming household appliances started with the 1975 Energy Policy and Conservation Act (EPCA). EPCA introduced test procedures, labeling, and energy efficiency standards related to energy use. The testing component required manufacturers to provide test results of all their marketed products, and the labeling component became the EnergyGuide program managed by the Federal Trade Commission. The Department of Energy was eventually tasked with implementing the minimum energy efficiency standards program.

EPCA was amended by the National Energy Conservation Policy Act (NECPA) of 1978. NECPA introduced the list of products initially covered by standards and provided more detail on the rulemaking process.

The first U.S. federal minimum efficiency standards were actually established under the National Appliance Energy Conservation Act (NAECA) of 1987. Among the affected products were clothes washers (effective January 1, 1988), dishwashers (effective January 1, 1988), room air conditioners (effective January 1, 1990), and refrigerators and freezers (effective January 1, 1990). These standards were primarily performance-based, though differing by product classes defined by technical characteristics in some cases, and focused on establishing a maximum energy use metric. However, for a small subset of products the standards were also technology-based, requiring certain features

[5] Manufacturers have often argued against some standard revisions on the basis that the new standard will force some technologies out of the market and reduce performance (AHAM 2015). It is explicitly written into law that the Department of Energy is not allowed to enact a new standard if it has the knowledge that the standard will reduce product availability and/or performance.

or options to be made available (e.g., a dry without heat option for dishwashers, and a cold water rinse option for clothes washers).

Beyond the initial round of NAECA standards, there was a requirement that each product go through two subsequent rulemakings to update the standards further if it was determined that the new standard level was technically feasible and economically justified. The1992 Energy Policy Act (the "1992 EPAct"), the 2005 Energy Policy Act (the "2005 EPAct"), and the Energy Independence and Security Act (EISA) of 2007 all affected the procedures and timelines of revised standards for appliance products. Recently introduced standards and revised standards established by NAECA were almost exclusively performance-based.

The process by which individual standards are set is a multi-year process. It takes three to five years from the initial steps, which consist of conducting a market and technology assessment (to determine if the standard for a given product will be revised), to the adoption of the final rule.[6] In addition, a new standard is generally not effective in the market for three years following the date when it is adopted by DOE. Finally, the timing of when different products are reviewed is generally set many years prior by overarching legislation, and the specific new rules are set multiple years prior to when they become effective.

In addition to the minimum efficiency regulation, appliances have also been subject to non-regulatory energy efficiency policy, the most prominent of which is the ENERGY STAR (ES) labeling program, a voluntary certification standard for energy intensive durables. In particular, the "1992 EPAct", further amended by the "2005 EPAct", established the ES program, which is administered jointly by the EPA and DOE. This program identifies and labels products with a high level of energy efficiency. The ES certification requirement is generally defined relative to the minimum standard level effective at any given point in time, though it can be updated more frequently. In particular, the 2005 EPAct specified that the ES criteria should be regularly updated, and in 2009 the EPA and DOE established that, other than in cases with appliances that had very long product life cycles, the ES criteria would be reviewed every three years, or more frequently if the market share of qualifying products exceeded 35 percent.

Our analysis focuses on six appliance categories: clothes washers, dishwashers, room air conditioners, freezers, compact refrigerators, and full-size refrigerators. Table 1 summarizes the history of the minimum efficiency standards and the ES certification requirements for these products. Each

[6]The time between Final Rule and compliance varies from product to product and it is determined by the legislation.

entry corresponds to the effective date that these regulations first came into effect or were revised. Our data cover the period 2001-2011, which was particularly active for both minimum and ES standards. Additional details about the standards for these appliance categories can be found in Appendix A. Not shown here is the depth and breadth of the coverage of these standards for the whole market of energy intensive durables. As of today, minimum standards cover more than 60 product categories. The DOE is still planning to expand this coverage and numerous revisions in existing standards are planned for the upcoming years. Similarly, for the ES program, which covers more than 70 product categories, further expansions and revisions are planned for the near future.

3. **Related Literature**

Performance-based minimum efficiency standards for appliances are instances of a combination of performance-based and attribute-based regulation (Ito and Sallee 2014), where the maximum amount of energy a given product can use is a function of a small set of product characteristics. Most appliance standards are set as a function of size or overall capacity such that larger appliances are allowed to use more energy. Other features such as the design of the appliance, for instance bottom-freezer versus top-freezer refrigerators, are also considered. The DOE's rationale in using attribute-based standards seems to ensure that the regulation does not restrict the choice set or distort the quality in the non-energy dimension. As shown by Ito and Sallee (2014), there is a formal efficiency argument to justify attribute-based standards. In the absence of a compliance trading mechanism, as in it is the case of appliance standards, attribute-based standards may improve efficiency by harmonizing compliance costs across products.

In practice, attribute-based standards are, however, likely to distort product quality. If the regulator has imperfect knowledge of manufacturers' costs, he will not be able to set optimal standards. As a result, standards will implicitly reward or penalize firms for providing certain attributes and will induce trade-offs between energy use and other attributes. Such trade-offs have been extensively documented in the passenger vehicle market (e.g., Knittel 2011; Klier and Linn 2012; Whitefoot, Fowlie, and Skerlos 2011). In the presence of imperfect information on the regulator side, they will invariably be synonym with an efficiency loss relative to the perfect information case.

Setting standards in the presence of a second market failure, imperfect competition, will also induce distortion in quality. The theoretical work looking at the interplay among market structure, quality provision, and minimum quality standard (MQS) is extensive and makes numerous, but sometimes contradictory predictions. As we discuss below, one robust prediction made by several models suggests that setting an MQS in the presence of imperfect competition may actually increase quality beyond the MQS and improve welfare relative to the unregulated case.

Spence (1975) was the first to provide the insight that market structure has important implications for the provision of product quality. He showed that a monopolist will provide a level of product quality that deviates from the social optimum due to the difference between the marginal versus the average consumer. The classic work of Mussa and Rosen (1978) extended this framework and considered heterogeneous preferences over a vertical quality attribute. They showed that the monopolist will underprovide quality to the consumers with a lower willingness-to-pay for quality, and charge increasingly higher margins for quality above the lowest provided. Besanko, Donnenfeld, and White (1988) then demonstrated the implications of an MQS (along with other regulatory mechanisms) on a monopoly market as initially described by Mussa and Rosen (1978) and showed that an MQS imposed in such a market would raise quality for the lowest consumer types, while leaving the effect on the quality provided to higher types of consumers unchanged. Leland (1979) and Shapiro (1983) considered the role of an MQS with a model of perfect competition in an imperfect information context. Both demonstrated that sellers in such a setting have an incentive to lower product quality in order to increase profits. However, Shapiro (1983) showed that, due to reputation effects, sellers also had an incentive to provide some high quality goods, but would charge a premium for these products. Both showed that an MQS could have positive impacts on the market including an increase in overall welfare in some cases (Leland 1979), but that some consumers could end up worse off as a result of the MQS and the regulation could result in the elimination of some products from the market that consumers would otherwise want to purchase.

While the monopoly and perfect competition cases are intuitive, the U.S. appliance market is best described as an oligopoly. A small number of manufacturers have large market shares and mergers and acquisitions in the last decade have substantially increased market concentration (Ashenfelter, Hosken, and Weinberg 2013). To illustrate, the Herfindal index for the appliance categories that we cover vary between 0.18 and 0.516.[7]

Ronnen (1991) introduced a model of oligopolistic competition (duopoly) to study MQS. In contrast to Leland (1979) and Shapiro (1983), he showed that in this type of market structure an MQS will result in higher quality products purchased by all consumers, and higher surplus to all

[7] In our data, the Herfindahl-Herschman indexes are the following: clothes washers 0.188, dishwashers 0.182, room air conditioners 0.277, refrigerators (compact and full-size) 0.177, and freezers 0.516.

consumers. The idea is that as the MQS causes the disparity in product quality differentiation to shrink, products become closer substitutes to each other, which causes prices to fall. High quality sellers raise the quality of the products they provide in order to alleviate this increased price competition. However, the range of quality in the market is still restricted by the MQS, which means that prices are still lower than in the unregulated case. This higher quality across the range of products, along with lower hedonic prices, means that consumers end up better off after the regulation.

Following Ronnen (1991) there is a rich literature, some of which we describe below, exploring the implication of an MQS with variations on the same basic model framework used by Ronnen (1991), which consists of a market characterized by the following: (i) a duopoly with a high quality seller and a low quality seller, (ii) generally identical single-product firms that compete in a two stage game in which qualities are chosen first and prices are chosen in the second stage with a solution being a subgame perfect Nash equilibrium, (iii) products that are differentiated both vertically (on some "quality" dimension), and horizontally using a variant on a Hotelling spatial style model, and (iv) usually the assumption that the market has full coverage (i.e., all consumers are catered to in the market).

One fundamental differentiating factor across variations of this model comes from assumptions regarding the cost structure for quality provision. Ronnen (1991) assumes quality-dependent fixed costs. Many other researchers have looked at implications of an MQS under different extensions of the basic model with this fixed cost assumption, including variations in timing of the quality competition phase (Constantatos and Perrakis 1998); more than two firms (Scarpa 1998; Pezzino 2010); Cournot quantity competition rather than Bertrand price competition (Valletti 2000; Pezzino 2010; Toshimitsu and Jinji 2007); and asymmetric costs across firms and endogenous quality ordering (Jinji and Toshimitsu 2004; Baliamoune-Lutz and Lutz 2010). Others have presented models with variable costs increasing with product quality. In particular, Crampes and Hollander (1995) assume variable costs of quality and show results consistent with Ronnen (1991). Others have further expanded the model with the variable cost of quality assumption to explore different settings, including endogenously set standards and the implications of an MQS on the capacity of firms to collude (Ecchia and Lambertini 1997); and asymmetric dominance across firms and the implication of an MQS on the distribution of market power (Ecchia and Lambertini 2001). Among all of the above work, the implications for an MQS have largely been consistent with the findings of Ronnen

(1991) and Crampes and Hollander (1995): an MQS results in increased quality and/or increased social welfare.[8]

Beyond minimum standard, the role of product labeling is also relevant to our setting, as appliances are subject to the ENERGY STAR (ES) label. The work that has explored the implications of product labeling has largely focused on variations of the model assuming some degree of uninformed consumers. Baltzer (2012) explores both an MQS and labels in a setting with asymmetric information and finds that labels are not necessarily the optimal policy when markets are imperfectly competitive. He finds that an MQS, when consumers are completely uninformed about product quality, results in higher welfare gains than a label. Buehler and Schuett (2014) also show that an MQS in a setting where a fraction of consumers observe quality results in higher quality and higher social welfare. This result, as in the case of Ronnen (1991) and Crampes and Hollander (1995), stems from the reduced product differentiation and resulting increased price competition. They show, on the other hand, that a labeling or certification program instead can result in firms having a tool to even further differentiate their products and can therefore increase profits. Only in the case when the share of informed consumers is very small is a certification program possibly more optimal relative to an MQS. While others have shown that labeling can be used to induce firms to increase quality that is depressed below the social optimum (Amacher, Koskela, and Ollikainen 2004), others show that with imperfect competition, information provision alone can actually cause welfare to fall if it induces too much product differentiation (Brouhle and Khanna 2007). Houde

[8]The only variations of the model in which the results are completely the opposite are the cases when the model is expanded to more than two firms, or the assumption is made that firms operate under Cournot competition rather than Bertrand competition. However, the results in various combinations of these two assumptions appear somewhat inconsistent. For instance, Scarpa (1998) showed that with three firms competing over prices, the introduction of an MQS decreases the maximum quality level provided and average quality consumed, and despite increases in consumer surplus, MQS can reduce overall welfare. However, he acknowledges that his results hinge largely on assuming no variable cost of quality, and he suggests that introducing these variable costs may change the main results. Pezzino (2010) showed that when there are three firms under Cournot competition, however, imposing an MQS has a positive effect on the average level of quality provided while Valletti (2000) shows that when two firms compete in a Cournot game, then an MQS unambiguously decreases welfare, although they do show that in aggregate consumers in each quality segment do benefit from the MQS. Jinji and Toshimitsu (2004) relax some assumptions of the model and show that the results of Ronnen (1991) in the Bertrand case, and Valletti (2000) in the Cournot case are largely robust. However, Napel and Oldehaver (2011) show that under Cournot competition, if the dynamic effects of standards impeding collusion are taken into account, then an MQS can result in overall welfare gains.

(2014a) showed that for the particular case of the ES certification, the program may be welfare improving, but most of the welfare gains accrued to firms that use the label to price discriminate.

In sum, the main insight from the theoretical literature is that setting standards in imperfectly competitive markets increases competition in the regulated dimension of the product space. Firms then have an incentive to distort quality to further differentiate their products and relax competition. This insight guides the second part of our empirical work in Section 7 where we investigate whether energy efficiency standards have an impact on quality beyond improvement in energy use.

4. Measuring Quality: A Revealed Preference Approach

The micro-foundation of our quality index is a discrete choice model in a market with differentiated products. Each consumer i chooses a product that maximizes his utility and values a product j as follows:

$$U_{ij} = \gamma_j - \eta(y_i - p_j) + \epsilon_{ij} \tag{1}$$

where p_j is the price of product j, y_i is the income of consumer i, and η, the coefficient on price, corresponds to the marginal utility of income. The term γ_j represents an index of quality common to all consumers and ϵ_{ij} represents idiosyncratic preferences, which can be thought of as quality in the horizontal dimension. Our goal is to estimate γ_j, a price-adjusted measure of vertical quality. We face two challenges in this exercise. First, it requires exogenous variations in prices to identify η. Second, ϵ_{ij} is unobserved. The classic approach to addressing the latter problem is to assume a distribution for ϵ_{ij}, such as the generalized extreme value distribution, which gives rise to the well-known multinomial logit formula for the choice probabilities. As pointed out by Caplin and Nalebuff (1991) and later by Berry and Pakes (2007), modeling idiosyncratic preferences with an additive i.i.d. error term amounts to assuming tastes for products. Ackerberg and Rysman (2005) and Berry and Pakes (2007) discussed why this assumption is problematic, especially in the presence of large choice sets where product entry/exit is an important phenomenon. It notably implies that the dimension of the unobserved product space increases proportionally with the number of products in the choice set. As a result, choice models that include an i.i.d. additive error term cannot capture congestion in the product space. This has unintended consequences for the identification of the model parameters and measurement of welfare. For instance, Ackerberg and Rysman (2005) demonstrate how this assumption can result in an identification of the coefficient on price in a logit model using data with no price variation.

In the present application, we are concerned about the identification of the product-specific fixed effects γ_j, which are also affected by how we model idiosyncratic preferences. A failure to account for congestion effects in the product space implies that the effect of (unobserved) horizontal product differentiation[9] will necessarily be captured by the product-specific fixed effects. As a result, the vertical component of quality will tend to be overestimated.

There have been several solutions proposed to this problem. Berry and Pakes (2007) proposed the pure characteristic model. This approach completely eliminates the horizontal component of unobserved preferences (i.e., the additive consumer-product-specific error term), and only accounts for a one-dimensional unobserved vertical product characteristic. The pure characteristic model is thus the complete opposite of a model with tastes for products. In this model, the expansion of the choice set will automatically lead to product congestion. This is also a restrictive assumption. In practice, one could estimate a model with tastes for products and also the pure characteristic model to bound the estimates and welfare effects (see Song 2007). The problem is that the pure characteristic model is computationally challenging to estimate. The algorithm to compute the estimator is not guaranteed to converge and for applications with a large number of products it performs poorly.[10]

Ackerberg and Rysman (2005) proposed an alternative estimator that can be qualified as a hybrid solution. Their approach consists of maintaining the assumption of additive i.i.d. error terms, but also including the number of products directly in the utility function as a way to control for unobserved horizontal product differentiation. They provide various micro-foundations for such an approach consistent with utility maximization. In practice, their approach can be implemented as a generalization of a logit-based model, where not all parameters necessarily have a clear structural interpretation.

For our estimation, we develop an estimator similar in spirit to the approach proposed by Ackerberg and Rysman (2005). We use a nested logit model with a flexible representation of the outside option, which, as we show below, is observationally equivalent to adding the number of products within each nest into the utility function. Moreover, we also include as a control a measure of distance between products, an approach also proposed by Ackerberg and Rysman (2005). In our

[9]Ackerberg and Rysman (2005) labeled the assumption of additive i.i.d. error terms the "symmetric unobserved product differentiation" (SUPD) assumption.

[10]To our knowledge, Song (2007) is the only application that used the pure characteristic model. His model was computed for only 20 products. In our application, the estimation must be performed with choice sets that well exceed 500 products. We had little success estimating the model with such large choice sets.

preferred specification, this distance function consists of an l^2-norm computed using a subset of observed attributes. The rationale of using such a distance function is to capture horizontal product differentiation using information about the location of the products in the observed characteristic space, instead of solely relying on the number of products.

Our estimating equation is derived as follows. First, we define the function $f(j, J, X_j, \delta)$ that takes as arguments the number of products J and observed characteristics X. In a nested logit model, the function could be simply defined as the log of the number of products in each nest g, i.e., $f(j, J, X_j, \delta) = \delta ln(J_{j|g})$, where δ is an unknown parameter and $f(j, J, X_j, \delta)$ enters the alternative-specific utility (Equation 1) additively. We can include this term and rearrange the choice probabilities using the Berry (1994) inversion to obtain:

$$ln(\sigma_j) - ln(\sigma_0) = \gamma_j - \eta p_j + \sigma\delta ln(J_{j|g}) + (1 - \sigma)ln(\sigma_{j|g}) \tag{2}$$

where σ_j and σ_0 are the market share of product j and the outside option, respectively. The term $\sigma_{j|g}$ is the market share of nest g to which product j belongs. Note that $\sigma_{j|g} = J_{j|g}/J$. Therefore, we can rearrange $\sigma\delta ln(J_{j|g}) + (1 - \sigma)ln(\sigma_{j|g})$ and write $(1 - \sigma + \sigma\delta)ln(J_{j|g}) - (1 - \sigma)ln(J)$. We define the parameter $\beta = 1 - \sigma + \sigma\delta$ and end up with the expression $\beta ln(J_{j|g}) - (1 - \sigma)ln(J)$. With panel data, adding time fixed effects allows us to flexibly model the outside option, which will then capture both $ln(\sigma_0)$ and $(1 - \sigma)ln(J)$. Therefore, only β can be identified. As alluded to above, the nested logit is therefore consistent with a particular model of product congestion. One well-known limitation of the nested logit is that the nest structure must be predefined based on ad hoc assumptions. Our strategy in addressing this issue is to use a data-driven approach to define the nest structure. We perform a k-means clustering analysis (Hastie, Tibshirani, and Friedman 2009) using a subset of important observed attributes that we observe for each appliance category. This procedure assigns each product to a cluster defined by a notion of distance between products and the clusters identified correspond to the nest structure of the choice model. Additional details can be found in Appendix B. In the results section, we also discuss sensitivity analysis on the nest structure.

Our preferred estimator will consider the more general function $f(j, J, X_j, \delta) = \delta ln(J_{j|g}) + \alpha R_{j|g}$, where $R_{j|g}$ is an l^2-norm that measures the distance across K attributes between product j and all other products located in nest g to which j belongs:

$$R_{j|g} = \sqrt{\sum_{k=1}^{K} \left(\frac{X_{jk} - \overline{X}_{gk}}{\sigma(X_{gk})} \right)^2},$$

where $\overline{X}_{gk}$ is the average value of attribute k within nest g and $\sigma(X_{gk})$ is the standard deviation. We rely on the l^2-norm as a measure of distance because it has a direct link with the k-means clustering analysis (see Appendix B) that we use to define the nests.[11] Formally, our function $R_{j|g}$ measures the distance from the center of each cluster identified by the clustering analysis.

5. **Data and Estimation Strategy**

The primary data used for the estimation are point of sales data provided by the NPD Group, a market research company. Each observation in the NPD data consists of the monthly total sales and total revenue for a particular appliance model. The data were collected from a large sample of retailers[12] during the 2001-2011 period and were aggregated at the national level. Each appliance model is identified by a unique manufacturer model number and is matched with basic attribute information, which varied depending on the appliance.

The definition of a product in our data is very disaggregated. For some appliance categories (e.g., room air conditioners (ACs), dishwashers, and clothes washers), the manufacturer model number corresponds to the actual product that consumers bought in stores. For other categories (refrigerators), a product is usually distinguished up to its color and a few options. In the refrigerator market, manufacturers commonly offer product lines with three to five different colors. Detailed refrigerator features such as ice-maker are observed.

The data were matched with various publicly available data sources to recover energy use and energy efficiency ratings. In particular, all appliances were matched to energy use data provided by the Federal Trade Commission (FTC). The FTC provides yearly model-specific energy consumption data, which reflect the information provided to consumers on the EnergyGuide label posted on

[11]Alternatively, Ackerberg and Rysman (2005) propose to define R_j as follows:

$$R_j = \sum_{k=1}^{J} \phi\left((X_j - X_k) * (cov^{-1}(X)) * (X_j - X_k)\right)$$

where ϕ is the pdf of the standard normal distribution. Ackerberg and Rysman (2005) describe the function R_j as a weighted sum of products with similar characteristics, but we find it more intuitive to label it a distance function. As the difference $X_j - X_k$ gets close to zero, the value of R_j increases.

[12]The number of retailers sampled by the NPD Group varies from year to year. NPD provided yearly estimates of the coverage of the U.S. appliance market; this varies from about 25% to 80% for all appliances except for room air conditioners, which had a lower market coverage (25% to 38%). In all cases the NPD data tended to cover less of the market in earlier years of our dataset, and the market coverage increased steadily over the sample period.

appliances prior to purchase. In addition, we matched the clothes washer data to energy usage data provided by the ENERGY STAR program as well as the California Energy Commission (CEC) in order to obtain data on the energy use metric specific to that used by the Department of Energy (DOE) in setting minimum efficiency standards.

In our analysis we take advantage of one additional source of rich data we have compiled for clothes washers specifically. We complemented the attribute information provided by NPD with information provided in the manufacturer-provided users' manuals.[13] From these users' manuals, we obtained detailed model-specific attribute information. In collecting the data, we were very careful in identifying the various nomenclature used by manufacturers to describe a technology and track the incidence of each particular technology over the sample period. We also collected information about the use of trademarks in describing features. Finally, we consulted with appliance experts to distinguish energy efficiency related characteristics from others. We first use this fine-grained attribute information to validate the NPD attribute data, which we found to be reliable. More importantly, we use this information to investigate how standards impact different product features and how the evolution of specific features correlates with overall quality.

5.1. **Estimation Strategy**

Our data consist of monthly revenue and sales by appliance model. Average monthly price can then be computed by simply dividing revenue by sales. Although we observe large variations in the average monthly price for each appliance model, this variation is likely to be correlated with product-time specific unobservables. Consider the effect of inventory and stock-out, for instance. As a particular appliance model arrives at the end of its product life, retail stores are more likely to offer deep discounts to clear inventories, which will lead to stock-outs. We will thus observe low national sales (due to stock-outs in regional stores) together with low prices, which suggests an upward sloping demand curve. In sum, the level of aggregation of the data and the presence of dynamic pricing strategies in the appliance market pose important challenges for the estimation of the coefficient on price.

To account for time varying unobservables correlated with a product life cycle, we add product age as a control variable. We measure age in months starting with the first month where we observe

[13]We collected the users' manuals from various online sources and extracted the content from the pdf documents using help from our valuable research assistants.

a sale in the panel data. We allow the effect of product age to be as flexible as possible by defining a dummy for each month of age.

For the coefficient on price, we calibrate the parameter $\eta = \overline{\eta}$ and conduct extensive sensitivity tests with respect to its value. We fix $\overline{\eta}$ so that the long-term price elasticity is approximately equal to -2.25, which is close to the value that Houde (2014b) found for the U.S. refrigerator market.[14] Our decision to calibrate the coefficient on price was based on pragmatism and transparency. By ruling out the estimation of η, our estimation is computationally simple and easy to replicate. Through sensitivity tests, we provide bounds on the value of the quality index and show that the choice of η does have an impact on the results, but not on our main conclusions.[15] This modeling decision is therefore rather inconsequential for the present study.

Our preferred specification is:

$$ln(\sigma_{jat}) = \gamma_j - \overline{\eta} p_{jt} + \beta ln(J_{j|g,t}) + \alpha R_{j|g,t} + \nu_a + \mu_t + \xi_{jt}, \tag{3}$$

where index a represents product age, which is simply the number of months since product j has been on the market, and t is a subscript for months of sample. The controls ν_a and μ_t are therefore month-of-age and month-of-sample fixed effects, respectively. We estimate the above model with OLS separately for each appliance category. Standard errors are clustered at the appliance model level.

In Equation 3, the quantities of interest are the product fixed effects: γ_j, which we use to construct a price-adjusted quality index. Our preferred index consists of a monthly model-weighted average of the product fixed effects:

$$QI_t = \sum_{j}^{J} D_{jt} \gamma_j \tag{4}$$

where D_{jt} is a dummy variable that takes a value of one if product j is on the market in month t, and zero otherwise. The index is model-weighted instead of sales-weighted because we want to show how the quality offered by manufacturers responds to standards and exclude demand substitution effects. The time variation in quality displayed by the index will thus come solely from appliance models exiting and entering the market every month, not from sales.

[14]We set the value of η to -0.0045 for all appliance categories. For an average price of \$500, this corresponds to an own-price elasticity of approximately -2.25 according to the logit model.

[15]In interpreting our results, we focus on the sign of the welfare effects, not the precise dollar amount.

Finally, note that we estimate the product fixed effects γ_j for a sub-sample, J, corresponding to the most popular appliance models in each appliance category. For full-size refrigerators, we consider the models that are responsible for 80% of the total sales in a given year, which yields 1,183 unique refrigerator models. For the other five appliance categories, we use the models that are responsible for 90% of the total sales.[16] Using this criterion, but the number of models in the sample for clothes washers, dishwashers, room air conditioners,[17] freezers, and compact refrigerators are: 708, 800, 650, 210 and 265, respectively.

6. Results: The Evolution of Quality in the Appliance Market

Figure 1 shows the evolution of the price-adjusted quality index (Equation 4) computed for the six appliance categories. The 95% confidence intervals are also shown and obtained via the delta method using the full covariance matrix of the product fixed effects estimated with Equation 3. The vertical lines identify the effective dates of the revised minimum energy efficiency standards (MES) and/or ENERGY STAR standards (ES).

Graphically, there are two important stylized facts that emerge. First, some quality indexes display a strong upward trend and those trends coincide with revisions in minimum and/or ENERGY STAR standards. Second, for several events, we observe a sharp change in quality exactly at the time that the new minimum and/or ENERGY STAR standards became effective. This suggests that there was a high rate of product turnover induced by the revised standards and manufacturers were able to coordinate their product line decisions to meet the standards. Figure 2 confirms this. We observe that the large changes in quality indeed coincide with large reductions in average product age—as old product lines exit the market and are replaced by new products, the average age of products present in the market drops.

Figure 1 also shows that apart from the 2007 revision in the ES standard for dishwashers, all the large discrete changes in quality at the time of revisions consist of an increase. This means that products that entered the market to meet the more stringent standards were of higher quality, on average, relative to the previous generations of products.

[16]We experimented with this criterion, using a lower or higher threshold did not have an important effect on the results.

[17]We distinguish between air conditioners with and without reverse cycle technology. In our main analysis, we present only results for air conditioners without reverse cycle. Additional results for models with reverse cycle can be found in Appendix E.

6.1. **Sensitivity Tests**

The above results are robust to various specifications and modeling assumptions. In this section, we only present results for clothes washers. Sensitivity tests for the other five appliance categories yield similar conclusions and can be found in Appendix C. Panel A of Figure 3 shows the impact of changing the value of the coefficient on price. To bound our estimates of quality, we re-estimated Equation 3 using a value of η that corresponds to an own-price elasticity close to zero (-0.25) and a value of η that corresponds to a much higher elasticity (-3.75). In our preferred specification, we set η such that the own-price elasticity is approximately -2.25. For a value of η close to zero, the overall pattern in the evolution of the quality index remains the same, but the upward trend is less pronounced. To understand this result, we need to also consider the evolution of prices. Panel A displays the model-weighted average price of all models on the market at a particular point in time.[18] We observe a sharp decrease in prices after the 2004 revision followed by a steady increase. Setting η close to zero implies that we have a weak adjustment for price in the estimation of quality. The quality index will thus capture the effect of a change in average price. If prices increase on average, this will bias the quality index downward. This is exactly what we observe. On the other hand, if we set η too high, the bias will be in the other direction. In the present case, a larger value of η has a small impact and leads to a slightly more pronounced increase in quality following the 2007 revision in the minimum standard (second vertical line). Overall, the calibration of the price coefficient does have an impact on the quality index, but the overall patterns are robust for a wide range of values. In Appendix C, we show similar results for the other appliance categories. In sum, the calibration of the price coefficient is rather inconsequential for our main qualitative conclusions.

Product age should be strongly correlated with a number of unobservables confounded with the vertical component of quality, such as product placement strategies (online and brick-and-mortar), inventories, product reviews, and sales agent strategies. Our flexible controls for product age aim to rule out these effects. Panel B shows that product age has an important impact on the slope of the quality index, but it has an even more important effect on the magnitude of the discrete change in quality at the 2004 revision. Without the controls for product age, we observe a sharp and large increase in quality because a large number of new models entered the market exactly the month following the standard change (Figure 2). There is then a strong first-month-of-age

[18] To compute this average price, we expressed all prices in 2011 $ and first computed the average price over the entire lifetime for each model in the sample. The average price that we report is the model-weighted average of the average lifetime prices.

effect[19] that—if left uncounted for—can be confounded by a large increase in quality. For all six appliance categories, we found that controlling for product age has the larger impact on the results. Nevertheless, our main conclusion that quality has tended to increase and is correlated with the timing of revision in standards is robust to this modeling assumption.

Panel C compares quality indexes that rely on different specifications to control for unobserved horizontal product differentiation. At one extreme, we do not include any controls and use a simple logit model. At the other extreme, we experiment with a nested logit where we substantially increase the number of nests. In our preferred specification, the k-means clustering analysis suggests that we rely on 5 clusters (nests) (Appendix B). For this sensitivity test, we consider a model with 28 nests. Overall, the nest structure has little impact on the quality index. We also found similar results when experimenting with the distance function. We believe that the fact that our product definition is very disaggregated combined with the fact that consumer preferences might be defined primarily over the main attributes play an important role here. Our results suggest that the product fixed effects do well in capturing all the time invariant attributes of a product and idiosyncratic consumer preferences are tightly distributed around these fixed effects.

Finally, Panel D compares the model-weighted index to the sales-weighted index. Because our primary goal is to understand how firms respond to more stringent standards, moments and/or quantiles of the model-weighted distribution of quality in the whole choice set are the most relevant.[20] A sales-weighted index would be more relevant to understand how consumers experienced quality change during that period, and thus it informs us about the welfare gain due to an increase in quality offered. In the present case, the sales-weighted quality index also increases over time, and the size of the increase during the whole sample period is larger relative to the model-weighted index, suggesting that any improvement in quality offered will be magnified by the increase in demand for those products. Again, this pattern holds in the other appliance categories.

6.2. **Quantifying the Change in Quality**

We propose four approaches to measure the change in quality at the time of a standard change: a first-difference estimator, a differences-in-difference estimator, a first-difference estimator with matching, and a differences-in-difference estimator with matching. For all four estimators, the first difference consists of comparing products that exited the market at the time of a standard revision

[19]This effect can readily be seen by inspecting the month-of-age fixed effects in the regression results.

[20]For the analysis, we focus on the mean, but we could easily consider the median or other quantiles.

to products that entered the market during the same period. In particular, for each standard revision in a given appliance category, we compare products that exited the year prior or the year of the standard change to the products that entered the market the year prior or the year of the revision. We focus on this sub-sample of products as an attempt to capture the marginal effect of the regulation. Our rationale is that products that exited just the year before or the year of a standard change may have done so because they did not comply with the new standard, and products that entered at this time were meant to replace these noncompliant products.[21] In Appendix D, we show that the above criteria used to identify products that are likely to be marginal to the regulation perform well. For instance, most clothes washer models that exited the years of the 2004 or 2007 revisions in the minimum standard do not meet the more stringent standard and entering models tend to bunch or exceed the new standard. We found similar patterns for other appliance categories.

For revision events where the minimum and ES standards were both revised simultaneously, we quantify the change in quality for ES certified products[22] and products that were never certified.[23] The goal here is to distinguish the effects of the two different types of regulations. It also shows how tightening energy efficiency standards for products at the lower or higher end of the energy efficiency distribution differs.

We exclude from the analysis regulatory events that consist of the expansion of the ES certification to new appliance categories. This includes the 2003 ES standards for compact refrigerators and freezers, the 2003 ES standard for room air conditioners without louvered sides, and the 2005 ES standard for room air conditioners with reverse cycle. Our focus is thus exclusively on regulatory

[21]We define noncompliance using the year prior to or the year of the revision because although manufacturers have the ability to time their inventory decisions to correspond with a standard change, several noncompliant products stay on the market after a new standard is enacted. The same is true for newly compliant products, which are mostly introduced after a standard change, but some appear in the market in the months preceding a revision effective date.

[22]Products that were certified, but lost the ES certification following a revision in the ES requirement are classified as being ES certified in the regression analysis. Note that these models correspond exactly to the products that we consider marginal to the regulation.

[23]In the appliance market, it can be argued that manufacturers' costs of certifying a given product are almost zero if that product meets the ES certification requirement. Products that were never ES certified thus correspond to products that consume too much electricity to earn the certification, but still meet the minimum standard.

changes that consisted of tightening of an existing standard. These are the events for which we can confidently identify compliant and noncompliant models.[24]

For the differences-in-difference (DiD) estimators, the goal of the second difference is to rule out temporal shocks and trends affecting the whole appliance market. We are particularly concerned by the effect of a long-term declining trend in appliance prices documented in Spurlock (2014). Moreover, whether the increase in quality in several appliance categories displayed in Figure 1 is entirely due to standards or reflects a secular trend for the whole appliance market is also unclear. We implement the DiD by estimating the change in quality associated with revision events in all six appliance categories with a single regression model that includes year-of-sample fixed effects. In particular, the model that we estimate is:

$$
\begin{aligned}
(5) \qquad Quality_{jtk} = & \sum_{s}^{S} \beta^{1s} Comp_j^s \times Time_t^s \times ES_j + \sum_{s}^{S} \beta^{2s} NoComp_j^s \times Time_t^s \times ES_j \\
& + \sum_{s}^{S} \beta^{3s} Comp_j^s \times Time_t^s \times NoES_j + \sum_{s}^{S} \beta^{4s} NoComp_j^s \times Time_t^s \times NoES_j \\
& + \alpha_t + \alpha_k + \epsilon_{jtk},
\end{aligned}
$$

where $Comp_j^s$ is a 1-0 dummy variable that takes a value of one if product j in the appliance category k is compliant with a new regulation s, and $NoComp_j^s$ is defined similarly, but for noncompliant products. Note that in this specification, the set of all regulation events across the six appliance categories is S, and each event is indexed by s. The variable $Time_t^s$ is a 1-0 dummy variable that takes a value one if the year t is before, during or after the revision event s. The dummy variables ES_j and $NoES_j$ identify ES-certified and non-ES-certified products, respectively. Finally, α_t represents year-of-sample fixed effects and α_k represents appliance category fixed effects. We implement the first-difference estimator with exactly the same regression model, except that we omit the year-of-sample fixed effects. In Equation 5, the coefficient β^{1s} measures the mean quality in the years preceding, spanning and following the revision event s for all compliant products, i.e., newly introduced models, that are ES certified. The coefficient β^{2s} has a similar interpretation for noncompliant ES-certified models. The change in quality for a revision event s for ES-certified

[24]When the DOE and the EPA decide to expand the ES certification to new products, they usually consider the share of products that exceed the minimum standard. In this case, compliant models are thus products that may have been on the market for a long time and were not first introduced to meet an energy efficiency standard.

products is thus given by $\Delta Quality^{s}_{ES} = \beta^{1s} - \beta^{2s}$. For non-ES-certified products, the change in quality is given by $\Delta Quality^{s}_{NoES} = \beta^{3s} - \beta^{4s}$.

Note that the change in quality estimated by Equation 5 (and similarly for the first-difference estimator) compares newly introduced products to exiting products controlling for only one attribute: whether or not a product meets the ES certification. It thus captures the change in quality in all dimensions of the products, except for prices and the ES certification. Moreover, it captures not only the changing quality within product lines, but also the effect of the introduction (scrapping) of new (old) product lines, i.e. change in quality across product lines. To illustrate, consider the case of clothes washers. For the 2004 revision, the important technological innovation that allowed manufacturers to increase energy and water efficiency was to change the design from top-loading to front-loading. During that period, some manufacturers, especially Whirlpool, were also better at designing front-load washers given that they introduced the first front-load models a few years prior to the 2004 revision. Following the 2004 revision, the change in quality was then driven by two effects. First, there was a large increase in the number of product lines with front-load models. Indeed, as shown in Figure 5, there was a large increase in model share for front-load washers that started in 2004. Second, existing product lines of front-load models were improved to match the 2004 minimum and ES standards. As a result, for this particular revision event, the change in quality captures both the introduction of a new innovation (front-loading) and incremental improvement in this technology. The composition of the quality index with respect to technology, and more generally with respect to the type of product lines, is thus changing over time. The goal of our matching estimators is to account for composition effects and thus estimate quality changes only within product lines, i.e., to capture incremental improvement in technology. We do so by identifying product lines that span the revision period of each standard and control for product line fixed effects. In practice, we implement this estimator by adding product line fixed effects, α_l, to Equation 5. To illustrate, a product line for clothes washers consists of all products offered by the same manufacturer and that have the same brand, same door type (front-load versus top-load), same size, and same price point, where size and price are defined by a categorical variable that take three values. We use similar criteria for other appliance categories. We refer to this approach as a matching estimator because it comes as close as possible to comparing products that manufacturers introduced to replace noncompliant products.

From a policy perspective, whether it is more appropriate to focus on the estimators with or without matching is open to discussion. In the rulemaking analysis for minimum energy efficiency standards, historically consumer net benefits associated with more stringent standards were computed holding all attributes constant, and only the trade-off between the purchase price and energy savings was considered. This practice includes all the standards that came into effect during the time period of our analysis. The implicit assumption in this practice is that following a revision, consumers can substitute for products that are identical in design, except for their energy use and purchase price. Substitution effects between products of different design (or even brands) are thus implicitly ruled out. The matching estimators aim to rule out substitution effects. They are thus the most appropriate estimates to inform the rulemaking analyses under this set of assumptions. However, since 2015, standard rulemaking analyses have incorporated a consumer choice model, which models the degree to which consumers may substitute between products of different design (i.e. product class), price point, and energy use level following a standard revision. The non-matching estimators are thus also equally important to consider as they also capture change in quality across product lines, and better reflect current standard analysis practices.

Table 2 reports the results for all four estimators. Across all appliance categories, we focus on the revisions that occurred between 2002 and 2010 because the years 2001 and 2011 correspond to entry and exit dates, respectively, for all products in our panel.[25] Under this criterion, we have three revisions in minimum standards that impacted two appliance categories: clothes washers and dishwashers. Note that for clothes washers, but not for dishwashers, a revision in the ES requirement was also concordant with the revision in the minimum standard. Focusing on the first-difference estimator (Specification I), with the exception of ES-certified dishwashers in 2010, we observe an increase in quality—new products that just entered the market following a standard revision were then of higher overall quality relative to the models that were pushed out of the market by the more stringent standards. For revision events that exclusively targeted the ES requirement, we also observe increases in quality. The increases are statistically significant, except for ES-certified dishwashers affected by the January 2007 revision and the ES-certified refrigerators affected by the January 2004 revision. The magnitude of the increases is also economically significant. Indeed, given the logit-based micro-foundation of the index, a relative change in quality has a welfare interpretation. For a given change in quality, we can compute the equivalent or compensating change in price that would bring the same level of utility.[26] For instance, suppose that the estimated

[25]Under our criterion to identify products that are marginal to a regulation change, all products present in the year 2001 or the year 2011 would be classified as marginal for standard revision that occurred in those years.

[26]The demand model excludes income effects. The equivalent and compensating variations are therefore equal to each other.

change in quality is 1; the equivalent/compensating variation will then be $\$1/|\eta|$. For our preferred value of $|\eta| = 0.0045$, this thus corresponds to \$222.[27] In Appendix C, we show how the results from Table 2 translate into a money metric and present the results for different values of $|\eta|$. We present regression results (Table C.1) for a value of $|\eta|$ closer to zero (0.0005), i.e., for a high marginal utility of income, and find that the increase in quality at the time of revisions tends to be smaller, but the signs remain the same and the results remain statistically significant. For a low marginal utility of income ($|\eta|$=0.0075), the estimated changes in quality are larger (Table C.2). Whereas the estimated changes in quality are robust to the value of the marginal utility of income used for the calibration, the estimates translated into a money metric are much more sensitive (Table C.3). For this reason, we refrain from basing our main conclusions on the exact economic value of the changes in quality, but focus on the sign and the fact that they are economically significant for a wide range of values of $|\eta|$.

Comparing the price-adjusted quality index to the change in prices can inform the welfare sign of revision in a standard. If the overall vertical quality increases and price decrease, on average, we have an unambiguously improvement in welfare for consumers. Table 7 presents regression results for the same four estimators using the log of price[28] or the log of electricity consumption (kWh/y) as a dependent variable. The results correspond to the percentage change in price or electricity consumption at the time of the revisions. For the 2004 revision in minimum standard for clothes washers, we observe a large and statistically significant decrease in prices (Specification I, Table 7). For the 2004 revision for clothes washers, we also found an increase in overall quality, which suggests an improvement in consumer welfare.

Adding year-of-sample fixed effects (Specification II, diff-in-diff) tends to decrease the magnitude of the estimates. This is true for both quality (Table 2) and price (Table 7), but less so for energy use (Table 7). This implies that quality and prices were likely to be subject to unobserved shocks that affected the whole U.S. appliance market during this period.

[27] An alternative interpretation of the estimated changes in quality can also be done as follows. If we were to compare two products with exactly the same price and outside option that belongs to the same product nest, an increase of 0.5 point in quality, for instance, would translate in a sales ratio of $exp(0.5) =\approx 1.6$. In a logit-based model, the sales ratio is simply given by $exp\,(\Delta Quality)$. To see this, consider the ratio of the choice probabilities of two products, e.g., $P^{new}/P^{old} = exp(U^{new})/exp(U^{old})$. If the two products are exactly the same, except for the level of vertical quality, we have $P^{new}/P^{old} = exp\left(Quality^{new} - Quality^{old}\right) =$. If the difference in quality is 0.5 point, then $P^{new}/P^{old} = exp(0.5) \approx 1.6$.

[28] For each product, we use the average deflated price over the lifetime of the product.

Adding product line fixed effects (matching estimators, Specifications III and IV) decreases the magnitude of the estimates, in most cases. This is to be expected. As explained above, the matching estimators capture only change in quality within a product line, and rule out change in quality across product lines. Although smaller, several of the estimates remain economically and statistically significant. For prices, adding product line fixed effects has an important effect on the magnitude of the estimates (Specification III and IV, Table 7). Again the magnitude of the estimates is smaller. For revision events where prices decreased under Specifications I and II, the estimates are now non-statistically significant. The matching estimators thus suggest that the changes in prices of the new models relative to the noncompliant models were somewhat modest, although we observe large increases in overall quality.

Across the various standard revisions, the January 2007 revision in the minimum and ES standards for clothes washers and the January 2009 revision in the ES standard for this same appliance category both led to a particularly large increase in quality. These results hold for the four estimators. For these revisions, the changes in prices were small and not statistically significant, which suggests that these more stringent standards may have made consumers better off.

7. Energy Efficiency Adjusted Quality

The price-adjusted quality index captures all time invariant attributes of a product. It does not inform to what extent manufacturers trade offs energy use with other attributes. As discussed in Section 3, energy efficiency standards can induce an expansion or contraction of quality in the non-energy dimensions of the product space for two reasons. First, the standards are attribute-based and thus it may be easier for manufacturers to meet a certain energy efficiency requirement by adjusting a non-energy attribute, such as size, implicitly targeted by the regulation. Second, setting standards in imperfectly competitive markets creates incentives for firms to further differentiate their products, which may lead to expansion in quality. In this section, we use two approaches to measure the evolution of quality in the non-energy dimension. The first approach provides a lower bound of an energy efficiency-price-adjusted quality index. Under this approach, we assume that all consumers discount future energy costs using a discount rate on par with other investment opportunities. In the second approach, we let the data reveal how consumers value future energy costs and estimate the marginal willingness to pay for specific attributes, as in a standard demand estimation with differentiated products (Berry, Levinsohn, and Pakes 1995). We implement this latter estimator by correlating the price-adjusted quality index with a large number of attributes using a LASSO regression. Using our estimates of the willingness to pay for energy efficiency, we can then compute an energy efficiency-price-adjusted quality index. Both approaches suggest that more stringent standards may have led to an expansion of quality in the non-energy dimension.

7.1. **Lower Bound Approach**

For most products in our sample, we observe the electricity consumption per year that is reported by the manufacturers.[29] Note that this information, together with an estimate of the yearly operating costs, is reported for all appliance models sold in the marketplace with the mandatory energy label EnergyGuide. The extent to which consumers use that information and trade off future operating costs with purchase prices in the appliance market is unclear. Houde (2014b) found evidence that some consumers may rely on energy information when purchasing refrigerators, but most may simply dismiss this information. Davis and Metcalf (2014) provide survey data consistent with this finding. Older studies (see Train (1985) for a comprehensive review) also found evidence suggesting that consumers may steeply discount energy costs in the appliance market.

In the present application, we will derive a lower bound on a measure of energy-efficiency-price-adjusted quality by assuming that **all** consumers trade off future energy operating costs in a way that is consistent with the opportunity cost of money they may face. For instance, if consumers can earn an average return of 5% on their investments, we will assume that they trade off future energy operating costs of the appliance they purchase with this discount rate. Under this assumption and for a given estimate of the expected lifetime of a product, there is a direct link between the marginal utility of income and how much consumers should value a stream of future electricity costs. To see this, assume that consumers do not account for the effect of depreciation, and compute the lifetime electricity costs ($LC_{r,j}$) by summing and discounting the expected annual electricity costs ($C_{r,j}$) over the lifetime of the durable:

$$LC_{r,j} = \sum_{t=1}^{L} \rho^t C_{r,j} = \frac{1-\rho^L}{1-\rho} C_{r,j}, \tag{6}$$

where L is the lifetime of the durable, $\rho = 1/(1+r)$ is the discount factor, and r is the discount rate. Consistent with the fact that $|\eta|$ corresponds to the marginal utility of income, the coefficient for the sensitivity to annual energy costs in a demand model, say θ, can be expressed as a reduced

[29]For some products, this information is missing. We simply exclude these products from the current analysis.

form parameter corresponding to:

$$\theta = \eta \frac{1-\rho^L}{1-\rho}. \tag{7}$$

For a given value of r, η, L, and electricity price (p_e), we can thus compute a price and energy efficiency-adjusted quality index: $\xi_j = \gamma_j - \theta p_e kWh/y$, where γ_j is the price-adjusted quality index estimated for a given value of η. We consider that the index ξ_j is a lower bound of a measure of quality in the non-energy dimension because it assumes that all consumers consider energy costs and do so in a way that is consistent with a market interest rate. Arguably, not all consumers may consider energy costs, and even if they do so they might discount them at a steeper rate. By over-estimating the valuation of energy costs, ξ_j is underestimated and thus corresponds to a lower bound.

Figure 4 shows the energy efficiency-price-adjusted quality index for five appliance categories.[30] Relative to the price-adjusted quality index, this index displays a lower level of quality, but we nonetheless observe increasing trends. This suggests that quality in the non-energy dimension might have been increasing.

Table 4 presents regression results for the same four estimators considered above. Compared with results of the price-adjusted index (Table 2), we only find a sign change for the January 2004 revision of the minimum standard for clothes washers. This decrease in quality is, however, largely compensated by the large increase in quality for the 2007 revision. Why did the two revision events for this same appliance category have different effects on quality in the non-energy dimension? To understand this, it is important to consider the institutional details of these two revisions. In the early 2000s, the DOE announced to the manufacturers a plan to tighten substantially the minimum standard for clothes washers. Manufacturers were, however, concerned that they could not meet this standard at a reasonable cost. Following negotiations with manufacturers and trade groups, the DOE opted to implement the tightening of the minimum standard in two phases, where the major revision became effective in January 2004 and a smaller incremental change occurred in 2007.[31] Our results show that the 2004 standard was stringent enough to induce manufacturers to trade off energy efficiency with other dimensions of quality. That is, meeting the higher energy

[30] Room air conditioners were excluded from this analysis because we do not have enough information to compute annual energy use for this appliance category. We only observe the energy efficiency ratio, which cannot be readily converted into electricity use.

[31] This explains why we see a large decrease in energy consumption in 2004, but a more modest decrease in 2007 (Figure 2).

efficiency requirements came at a cost of a decrease in quality. We find these results particularly plausible given that there were numerous consumer complaints and class-action suits that were made regarding front-load washers produced during the 2001-2008 period.

For the other appliance categories, results from Table 4 suggest an increase in quality in the non-energy dimension at the time that the standards were revised, except for ES-certified dishwashers in 2010. Note that for dishwashers, there is a discrepancy between the graphical evidence of Figure 4 and the regression results for the January 2007 revision of the ES standard. According to the regression results, the quality of newly ES-compliant dishwasher models increased relative to noncompliant models exiting the market during the 2006-2007 period, but the index displays a sharp decrease in January 2007. This difference is due to the fact that several newly compliant models of high quality, as suggested by our estimates, entered the market a few months before the revised standard was enacted and several noncompliant models with low quality stayed on the market a few months after the revision effective date, which contributes to a sharp decrease graphically. The regression model, however, averages the quality of all the compliant and noncompliant models irrespective of the entry or exit month during the period 2006-2007.

Overall, even after making a generous adjustment for how consumers may value energy costs, we find evidence that quality in the non-energy dimension has increased in several instances. We now turn to our second approach and reach a similar conclusion.

7.2. **LASSO Approach**

In this second approach, we turn to estimation to quantify the valuation of energy costs θ. We only perform the estimation for clothes washers given that this is the only category for which we have an extensive list of product characteristics. From the users' manuals, we compiled a list of more than 45 different product characteristics describing the various technologies used by manufacturers over the sample period (see Table F.1, Appendix F).

For the estimation, we correlate the price-adjusted quality index of each product j with a vector of product characteristics X_j. In essence, we are performing the intermediate step of Berry, Levinsohn, and Pakes (1995)'s estimation procedure, where mean utilities are regressed on product characteristics. In our case, however, we are not concerned with the estimation of the price coefficient and possible correlation with price and unobserved characteristics, given that we ruled out the effect of price via calibration. Our challenge is to work with a high dimensional product space, where some characteristics might be collinear with others, and not all of them might be valued by consumers. Formally, we are dealing with a high dimensional sparse demand model. Recent advances in econometrics (Belloni, Chernozhukov, Hansen, and Kozbur 2014; Belloni, Chernozhukov, and Hansen 2014) have been proposed for computing and making inference with such models. We draw from this literature and especially from Gillen, Shum, and Moon (2014), who propose the LASSO-BLP algorithm. Again, the fact that we rely on calibration to account for the effect of price, greatly simplifies our estimation. Gillen, Shum, and Moon (2014) propose a three-step estimation procedure. In our case, we will simply run one LASSO-type penalized regression (Tibshirani 1996) where the goal is to find the elements of the whole vector of product characteristics X_j that best predict the price-adjusted quality index. Before turning to estimation, we show graphical evidence that the quality index for clothes washers is correlated with some salient attributes, the number of attributes, and trademarking.

7.2.1. ***Stylized Facts***

Panel A of Figure 5 shows the model share of front-load models over time. We see an increasing trend, which started around the time of the first revision in the minimum standard in January 2004 and tapered off toward the end of 2009, just before the revision in the ES standard. According to industry experts, the front-load design was crucial to achieve the important gain in energy efficiency observed in 2004 and was the main innovation used to meet the more stringent standard that became effective in 2004. We also observe that size (Panel A, gray dotted line), measured by tube capacity, has also been steadily increasing and the trend became more pronounced following the second revision in the minimum standard in January 2007. Note that the minimum and ES standards for clothes washers differ for compact and standard sized models and are set as a function of size within each of these product class categories. The expansion in tube capacity could then be partly the result of the new stringent standards that would have implicitly incentivized manufacturers to meet the energy efficiency standard with larger size models.

Panel B shows two proxies for motor performance, the spin speed and the number of cycles.[32] Spin speed has been steadily increasing, but the number of cycles first increased following the 2004 revision and then dropped sharply at the time of the 2007 revision. It thus appears that to meet the more stringent standards, the technology may have required a trade-offf with cycle technology and energy efficiency.

[32]The number of cycles refers to the number of separate cycle options a consumer can select (e.g., permanent press, regular, heavy duty, extra rinse). This variable is coded as a categorical variable that identifies a range instead of the precise number of cycles.

Panels C through F show the evolution of the number of features during the sample period. In all panels, we distinguish between front-load and top-load models. To count the number of features present on a given appliance model, we use a dummy variable that takes the value of one if a particular feature was listed in the user's manual and zero otherwise. In these data, a feature is defined as a particular technology that we track over the whole sample period. Panel C shows the average number of features on a given appliance model. We observe an increase in the number of features for both types of clothes washers, but a much larger increase for front-load. During this period, the front-load design was an important innovation. It is thus interesting to observe that innovation in the overall design was also accompanied by the addition of new features. In Panel C, we also observe a discrete increase in the number of features for top-load models at the 2007 revision. This means that the models introduced to comply with the new regulation were also subject to some innovation. Panels D and E distinguish whether these new features were directed toward achieving energy efficiency or to other dimensions. Panel D shows the evolution of energy efficiency-related features, which are features that have been identified by discussions with appliance experts as enabling higher energy efficiency performance (see Table F.1, Appendix F). We find that for both front-load and top-load models these features increased throughout the sample period with a sharp increase at the 2007 revision. Panel E shows non-energy efficiency-related features. For front-load models, we still observe a sharp increase in the average number of features at the 2007 revision. For top-load models, we also observe an increase in non-energy efficiency features, but of much smaller magnitude compared to the energy efficiency-related features. Altogether, this suggests that for front-load models, innovation at the time of the revision in standards was directed toward both energy efficiency and other dimensions of the product space. On the other hand, for top-load models, innovation was mostly directed toward energy efficiency.

Finally, Panel F provides additional evidence that manufacturers directed most of their innovation toward front-load models during that period, but that innovation was not purely related to energy efficiency. We observe that on average the number of trademarked features has increased for both types of models, but the rate of increase is much more pronounced for front-load relative to top-load models. For both types of models, we also observe a sharp increase in the year 2007. These trends in trademarking could suggest that manufacturers attempted to signal higher quality and differentiate their products beyond energy efficiency following the revisions in standards.

7.2.2. ***Estimation***

The above stylized facts show that the revisions in standards for clothes washers led to the introduction of new features. The overall increase in quality (Figure 1, Panel A) is thus correlated with an increase in energy efficiency, features and trademarks. The estimator that we propose aims to distinguish whether consumers value improvement in energy efficiency itself, or the various features that may or may not be related to energy efficiency.

The estimator consists of a LASSO-type penalized regression method, a well-known machine learning technique, which has only been used in economics recently. In essence, a LASSO regression minimizes the mean square error plus a penalty term that consists of the sum of the coefficients. The penalty term is an l^1-norm and thus sets to zero the coefficients that have a weak correlation with the dependent variable. The LASSO approach thus selects the most powerful predictors from a large set of potential explanatory variables. This technique is a data-driven selection method that can be applied in contexts where the number of regressors is large, possibly larger than the number of observations, and multi-colinearity is severe.

For the present application, we use the Post-LASSO method, which, as discussed by Belloni, Chernozhukov, and Hansen (2014), has the advantage of minimizing the bias toward zero that arises in the standard LASSO approach. It allows inference by recovering valid confidence intervals. We implement the Post-LASSO approach using the algorithm proposed by Friedman, Hastie, and Tibshirani (2009) and implemented in R. Additional details on our implementation can be found in Appendix F.

Tables 5 and 6 presents the estimation results. We estimated a separate LASSO regression for front-load and top-load models after experimenting with interaction terms.[33] The first important results are that the coefficients on energy use (kWh/y) and size (tube capacity) are both selected, are statistically significant and are of expected sign. The coefficients are also of similar magnitude for front-load and top-load models. Using the same assumption used in the precedent section to calibrate the valuation of energy costs, the coefficients on electricity consumption for front-load and top-load models correspond to an implicit discount rate of 32% and 62%, respectively. This suggests a substantial undervaluation of future energy costs compared to the 5% discount rate that we used above. In Tables 5 and 6, we compare the results from the Post-LASSO method with a simple OLS regression. The OLS estimates of the coefficient on electricity consumption come close

[33]We found that once we interact all features with a dummy that distinguishes front-load and top-load models, the dummy itself was not selected by the LASSO algorithm.

to the LASSO estimates. This means there is a strong correlation between price-adjusted quality and the measure of energy use reported by manufacturers, and this correlation is robust and likely not to be confounded by a correlation with other attributes.

Using the LASSO estimates, we compute a price-energy efficiency-adjusted quality index by predicting quality excluding the coefficient on electricity consumption. That is, we take the vector of estimated LASSO coefficients, $\hat{\beta}^{LASSO}$, set the coefficient on energy use to zero, and predict quality by computing $\hat{\xi}_j = \hat{\beta}^{LASSO}_{EE=0} X_j$. Figure 6 shows this predicted index. As a benchmark, we also show the price-adjusted quality index predicted by the LASSO approach, and the index computed in Section 5 obtained directly from estimating γ_j. Comparing these two price-adjusted indexes, the LASSO approach overpredicts the increase in quality at the 2004 revision, but it tends to under-predict the increase in quality in subsequent years relative to the estimated index. Note that the large jump in quality predicted by the LASSO approach captures the sharp decreases in energy use induced by the 2004 revision (Figure 2, Panel A). As shown by the sensitivity tests, the estimated index probably tends to underpredict this increase due to the month-of-age fixed effects. Once we remove the effect of electricity cost, the LASSO approach shows a steady increase in quality in the non-energy dimension over the whole sample period, where the increasing trend starts following the 2004 revision and increases further following the 2007 revision. In Tables 5 and 6, we quantify the change in this price-energy efficiency-adjusted quality index. We present only the first-difference estimators because we only performed the estimation for one appliance category. The regression results are consistent with the previous findings and suggest that for most of the revision events, quality of newly compliant models increases relative to noncompliant models.

An interesting finding from Figure 5 is that the average size of clothes washers has been increasing over time and was affected by the revisions in the minimum standards. In the LASSO regressions, the coefficients on size are relatively large and statistically significant for both front-load and top-load models, which suggests that consumers have a high willingness to pay for larger clothes washers. Therefore, how much of the increase in quality in the non-energy efficiency dimension can be attributed to manufacturers offering large models? Figure 6 shows the evolution of a quality index where the effect of both energy efficiency and size were removed.[34] We observe that the index remains constant until the 2007 revision, after which it starts decreasing. Figure 6 thus implies that the increase in features and trademarks in the post-2007 period was not a main driver of improvement in quality during this period.

[34]To compute this index, we proceeded as before. We predicted quality using the LASSO estimates where the coefficients on energy use and size were set to zero.

8. **Conclusion**

In this paper, we estimate the price-adjusted quality of products in the main appliance categories. We show that overall quality has increased or remain constant in the U.S. appliance market during the 2001-2011 period. More importantly, we observe large discrete changes in quality that coincide with revisions in energy efficiency standards, which include both minimum and energy efficiency standards. We present various estimators to quantify the changes in quality and price for appliance models that are the most likely to be marginal to the regulation. In most instances, we find economically and statistically significant increases in quality, but no statistically significant changes in prices. Together, these two effects suggest that more stringent standards may have been welfare improving for consumers. We also find evidence that the increase in overall quality is driven by improvement in energy efficiency over time, but not entirely. We show that quality in the non-energy dimensions has increased. We also document that manufacturers may have attempted to increase quality following revisions in standards by increasing features and trademark technologies.

The fact that energy efficiency standards could lead to improvement in quality beyond energy efficiency, increase product diversity, but also have little effect on prices challenges the common paradigm. These findings are, however, consistent with the presence of imperfect competition in this market. Several theoretical papers have demonstrated that imperfect competition could, in fact, be a rationale for quality standards. Our findings are only a preliminary step in understanding the role that imperfect competition combined with standards plays in the market for energy intensive durables.

References

Ackerberg, D. A., and M. Rysman (2005): “Unobserved Product Differentiation in Discrete-Choice Models: Estimating Price Elasticities and Welfare Effects,” *The RAND Journal of Economics*, 36(4), 771–788.

AHAM (2015): “AHAM Comments on DOE’s NOPR for Energy Conservation Standards for Residential Dishwashers; Docket No. EERE-2014-BT-STD-0021; RIN 1904-AD24,” Discussion paper.

Allcott, H., and D. Taubinsky (2013): “The Lightbulb Paradox: Evidence from Two Randomized Experiments,” Discussion paper, National Bureau of Economic Research.

Amacher, G. S., E. Koskela, and M. Ollikainen (2004): “Environmental Quality Competition and Eco-Labeling,” *Journal of Environmental Economics and Management*, 47(2), 284–306.

Ashenfelter, O. C., D. S. Hosken, and M. C. Weinberg (2013): "The Price Effects of a Large Merger of Manufacturers: A Case Study of Maytag-Whirlpool," *American Economic Journal: Economic Policy*, 5(1), 239–261.

Bajari, P., and C. L. Benkard (2005): "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach," *Journal of Political Economy*, 113(6), 1239–1276.

Baliamoune-Lutz, M., and S. Lutz (2010): "Pre-Emption, Predation, and Minimum Quality Standards," *International Economic Journal*, 24(1), 111–123.

Baltzer, K. (2012): "Standards vs. Labels with Imperfect Competition and Asymmetric Information," *Economics Letters*, 114(1), 61–63.

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "High-Dimensional Methods and Inference on Structural and Treatment Effects," *The Journal of Economic Perspectives*, 28(2), 29–50.

Belloni, A., V. Chernozhukov, C. Hansen, and D. Kozbur (2014): "Inference in High Dimensional Panel Models with an Application to Gun Control," *arXiv preprint arXiv:1411.6507*.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

——— (2004): "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 112(1), 68–105.

Berry, S., and A. Pakes (2007): "The Pure Characteristics Demand Model," *International Economic Review*, 48(4), 1193–1225.

Besanko, D., S. Donnenfeld, and L. J. White (1988): "The Multiproduct Firm, Quality Choice, and Regulation," *The Journal of Industrial Economics*, 36(4), 411–429.

Brouhle, K., and M. Khanna (2007): "Information and the Provision of Quality Differentiated Products," *Economic Inquiry*, 45(2), 377–394.

Brucal, A., and M. Roberts (2015): "Can Energy Efficiency Standards Reduce Prices and Improve Quality? Evidence from the US Clothes Washer Market," Discussion paper.

Buehler, B., and F. Schuett (2014): "Certification and Minimum Quality Standards When Some Consumers Are Uninformed," *European Economic Review*, 70, 493–511.

Caplin, A., and B. Nalebuff (1991): "Aggregation and imperfect competition: On the existence of equilibrium," *Econometrica: Journal of the Econometric Society*, pp. 25–59.

Carroll, S. L., and R. J. Gaston (1981): "Occupational Restrictions and the Quality of Service Received: Some Evidence," *Southern Economic Journal*, 47(4), 959–976.

——— (1983): "Occupational Licensing and the Quality of Service," *Law and Human Behavior*, 7(2-3), 139–146.

Chipty, T., and A. D. Witte (1997): "An Empirical Investigation of Firms' Responses to Minimum Standards Regulations," Discussion paper, National Bureau of Economic Research.

Constantatos, C., and S. Perrakis (1998): "Minimum Quality Standards, Entry, and the Timing of the Quality Decision," *Journal of Regulatory Economics*, 13(1), 47–58.

Crampes, C., and A. Hollander (1995): "Duopoly and Quality Standards," *European Economic Review*, 39(1), 71–82.

Dale, L., C. Antinori, M. McNeil, J. E. McMahon, and K. S. Fujita (2009): "Retrospective Evaluation of Appliance Price Trends," *Energy Policy*, 37(2), 597–605.

Davis, L. W., and G. E. Metcalf (2014): "Does Better Information Lead to Better Choices? Evidence from Energy-Efficiency Labels," Discussion paper, National Bureau of Economic Research.

Ecchia, G., and L. Lambertini (1997): "Minimum Quality Standards and Collusion," *The Journal of Industrial Economics*, 45(1), 101–113.

——— (2001): "Endogenous Timing and Quality Standards in a Vertically Differentiated Duopoly," *Recherches Économiques de Louvain/Louvain Economic Review*, 67(2), 119–130.

Feenstra, R. C., and J. Romalis (2014): "International Prices and Endogenous Quality," *The Quarterly Journal of Economics*, 129(2), 477–527.

Friedman, J., T. Hastie, and R. Tibshirani (2009): "Glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models," *R package version*, 1.

Gillen, B. J., M. Shum, and H. R. Moon (2014): "Demand Estimation with High-Dimensional Product Characteristics," in *Bayesian Model Comparison*, pp. 301–323. Emerald Group Publishing Limited.

Gormley, W. T. (1991): "State Regulations and the Availability of Child-Care Services," *Journal of Policy Analysis and Management*, 10(1), 78–95.

Hallak, J. C., and P. K. Schott (2011): "Estimating Cross-Country Differences in Product Quality," *The Quarterly Journal of Economics*, 126(1), 417–474.

Hastie, T., R. Tibshirani, and J. Friedman (2009): *The Elements of Statistical Learning*, vol. 2. Springer.

Hausman, J. A. (1979): "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables," *The Bell Journal of Economics*, 10(1), 33–54.

Hausman, J. A., and P. L. Joskow (1982): "Evaluating the Costs and Benefits of Appliance Efficiency Standards," *The American Economic Review*, 72(2), 220–225.

Houde, S. (2014a): "Bunching with the Stars: How Firms Respond to Environmental Certification," Discussion paper, University of Maryland.

Houde, S. (2014b): "How Consumers Respond to Environmental Certification and the Value of Energy Information," Discussion paper, National Bureau of Economic Research.

Ito, K., and J. M. Sallee (2014): "The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards," Discussion paper, National Bureau of Economic Research.

Jinji, N., and T. Toshimitsu (2004): "Minimum Quality Standards under Asymmetric Duopoly with Endogenous Quality Ordering: A Note," *Journal of Regulatory Economics*, 26(2), 189–199.

Klier, T., and J. Linn (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," *The American Economic Review*, 101(7), 3368–3399.

Leland, H. E. (1979): "Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards," *Journal of Political Economy*, 87(6), 1328–1346.

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

Napel, S., and G. Oldehaver (2011): "A Dynamic Perspective on Minimum Quality Standards under Cournot Competition," *Journal of Regulatory Economics*, 39(1), 29–49.

Parry, I. W., D. Evans, and W. E. Oates (2014): "Are Energy Efficiency Standards Justified?," *Journal of Environmental Economics and Management*, 67(2), 104–125.

Pezzino, M. (2010): "Minimum Quality Standards with More Than Two Firms under Cournot Competition," *The IUP Journal of Managerial Economics*, 8(3), 26–45.

Ronnen, U. (1991): "Minimum Quality Standards, Fixed Costs, and Competition," *The RAND Journal of Economics*, 22(4), 490–504.

Scarpa, C. (1998): "Minimum Quality Standards with More Than Two Firms," *International Journal of Industrial Organization*, 16(5), 665–676.

Shapiro, C. (1983): "Premiums for High Quality Products as Returns to Reputations," *The Quarterly Journal of Economics*, 98(4), 659–679.

Song, M. (2007): "Measuring Consumer Welfare in the CPU Market: An Application of the Pure-Characteristics Demand Model," *The RAND Journal of Economics*, 38(2), 429–446.

Spence, A. M. (1975): "Monopoly, Quality, and Regulation," *The Bell Journal of Economics*, 6(2), 417–429.

Spurlock, C. A. (2014): “Appliance Efficiency Standards and Price Discrimination,” Discussion paper, Lawrence Berkeley National Laboratory.

Tibshirani, R. (1996): “Regression Shrinkage and Selection via the Lasso,” *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), 267–288.

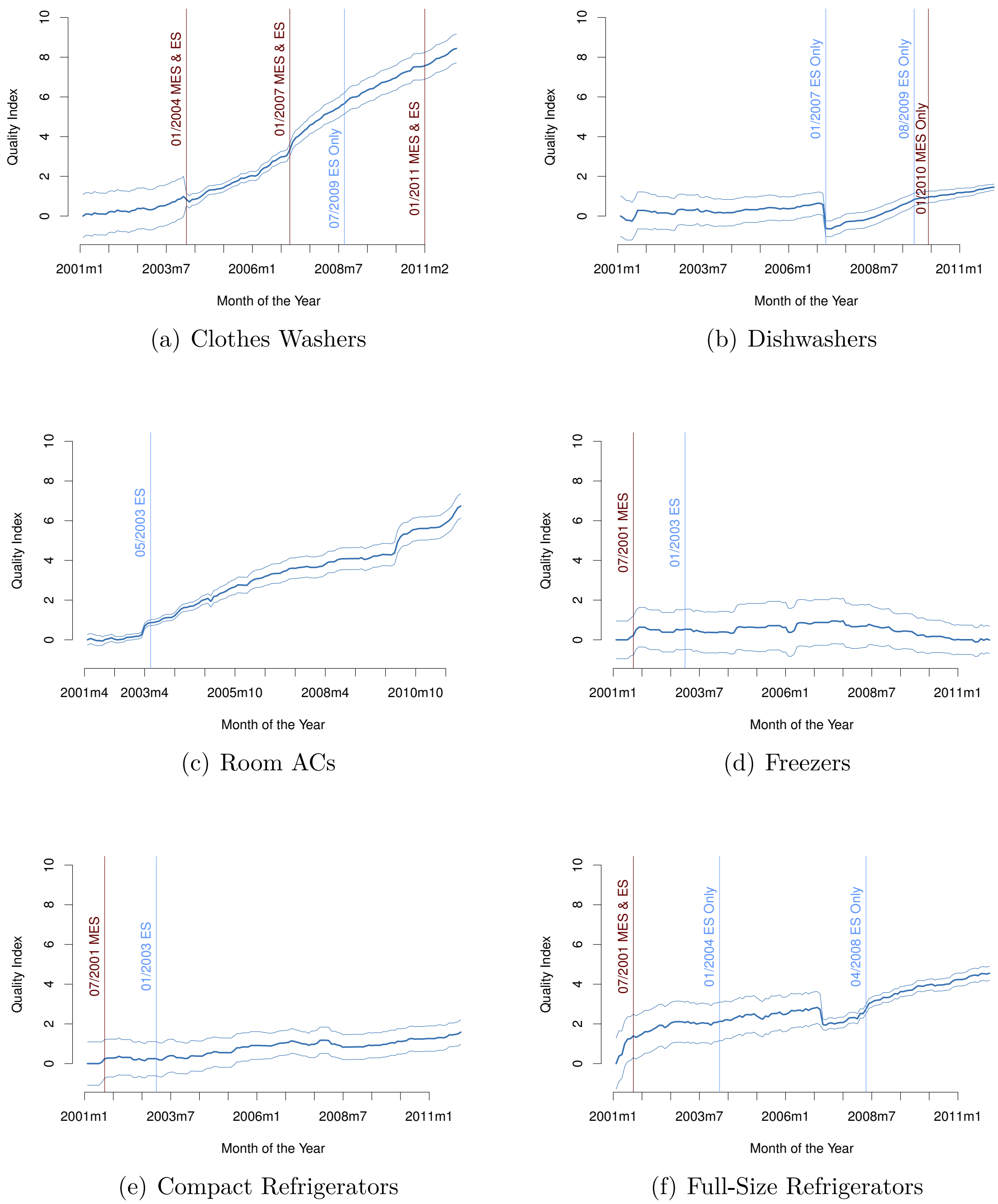
Toshimitsu, T., and N. Jinji (2007): “Quality Differentiation, Welfare, and the Mode of Competition in a Vertically Differentiated Product Market: A Note,” *Japanese Economic Review*, 58(3), 407–416.

Train, K. (1985): “Discount Rates in Consumers’ Energy-Related Decisions: A Review of the Literature,” *Energy*, 10(12), 1243–1253.

Valletti, T. M. (2000): “Minimum Quality Standards under Cournot Competition,” *Journal of Regulatory Economics*, 18(3), 235–245.

Whitefoot, K., M. Fowlie, and S. Skerlos (2011): “Product Design Response to Industrial Policy: Evaluating Fuel Economy Standards Using an Engineering Model of Endogenous Product Design,” *Energy Institute at Haas Working Paper No. WP-214*.

Wiggins, S. N. (1981): “Product Quality Regulation and New Drug Introductions: Some New Evidence from the 1970s,” *The Review of Economics and Statistics*, 63(4), 615–619.



(a) Clothes Washers

(b) Dishwashers

(c) Room ACs

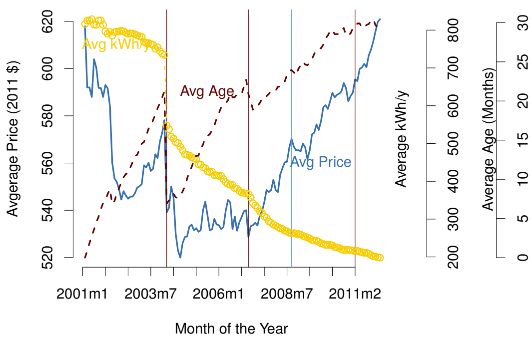
(d) Freezers

(e) Compact Refrigerators

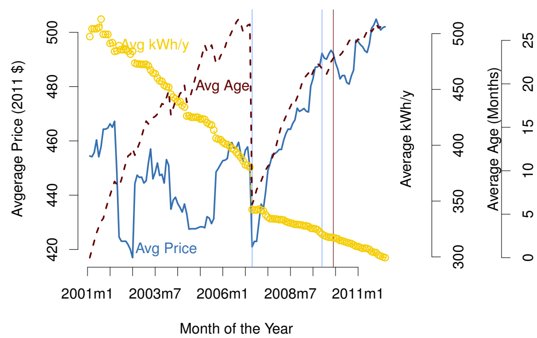
(f) Full-Size Refrigerators

FIGURE 1. The Evolution of Quality in the Appliance Market

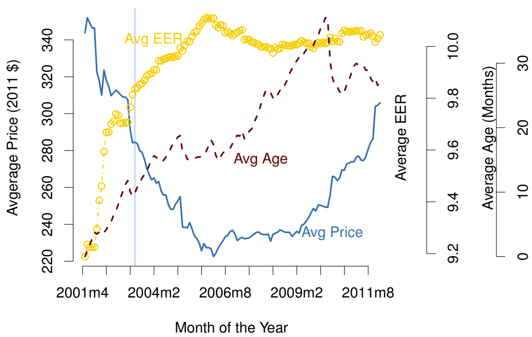
.



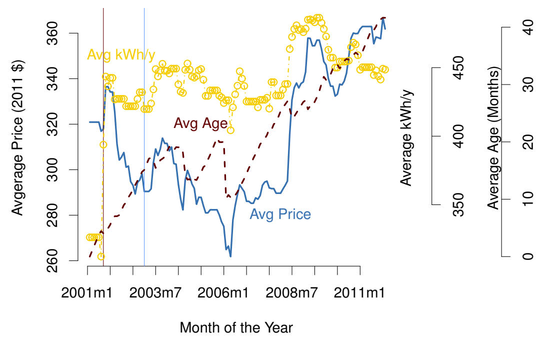
(a) Clothes Washers



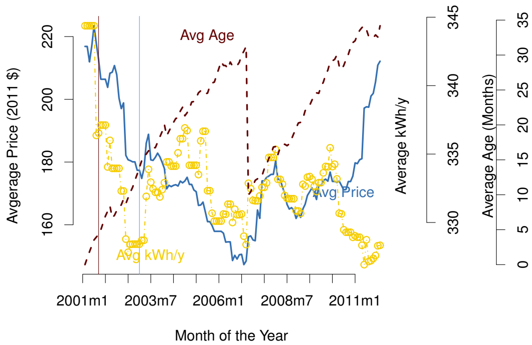
(b) Dishwashers



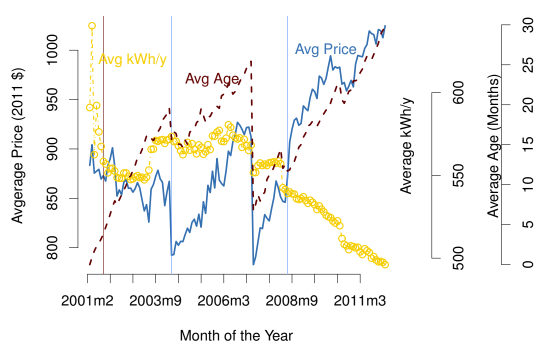
(c) Room ACs



(d) Freezers



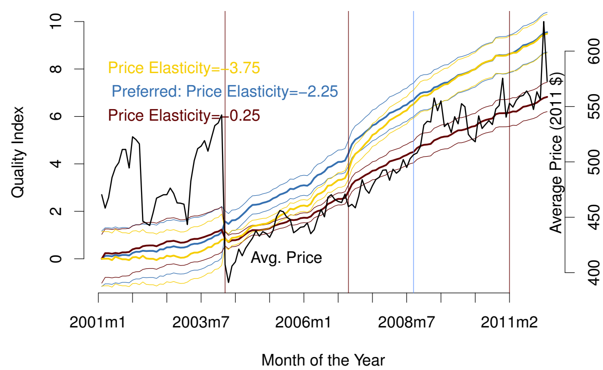
(e) Compact Refrigerators



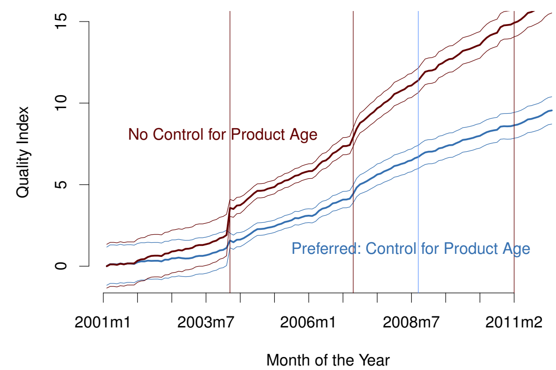
(f) Full-Size Refrigerators

FIGURE 2. The Evolution of Price, Energy Consumption, and Product Age in the Appliance Market

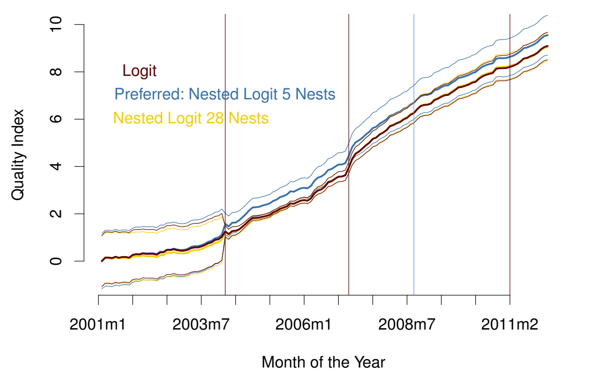
.



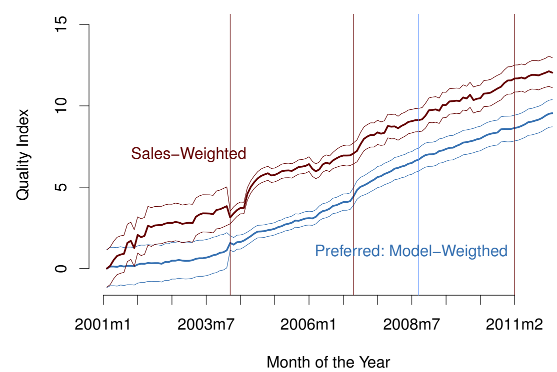
(a) Marginal Utility of Income



(b) Product Age



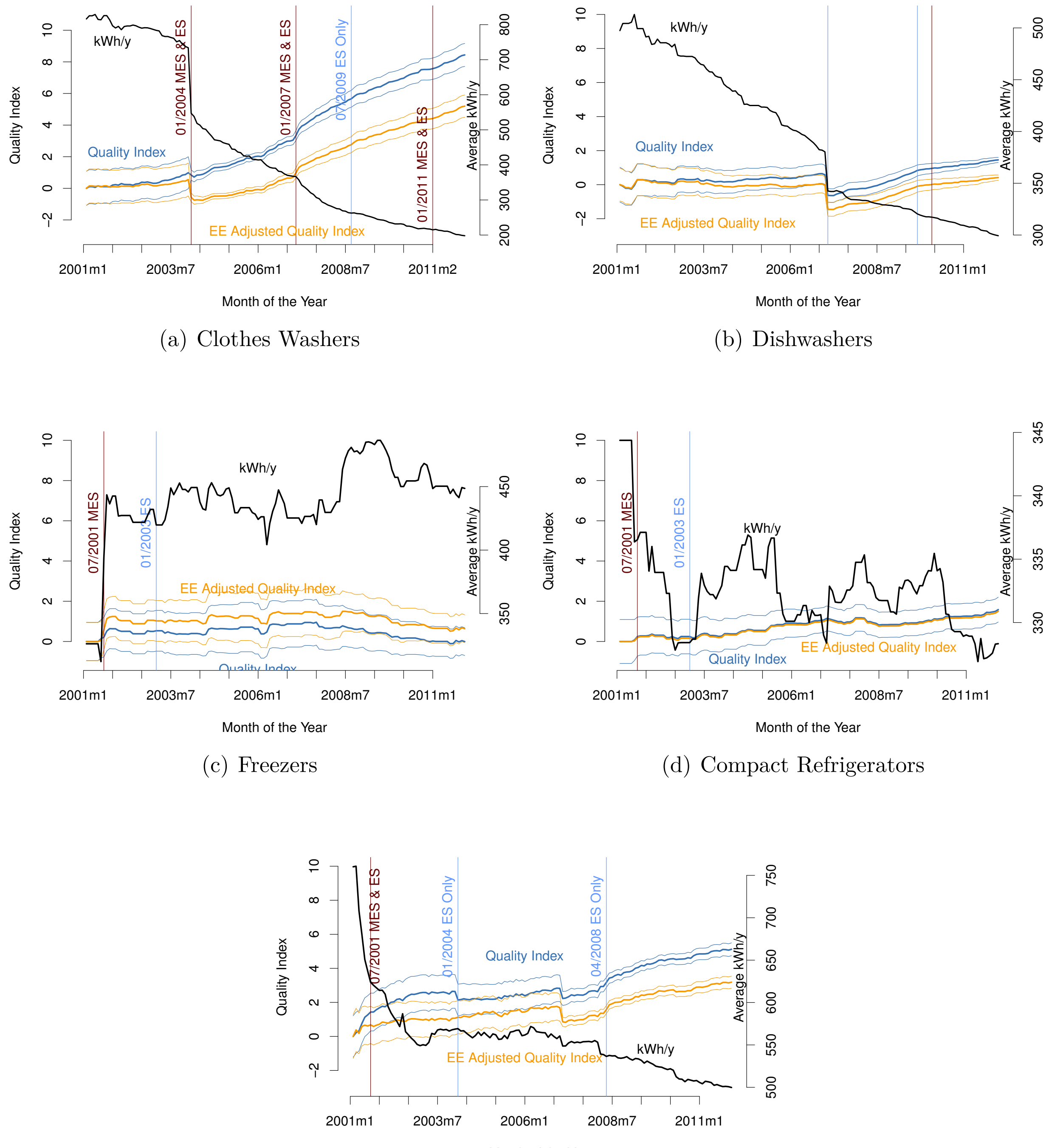
(c) Nested Logit Structure



(d) Sales vs. Model Weighted

FIGURE 3. Sensitivity Tests: Clothes Washers

.



(a) Clothes Washers

(b) Dishwashers

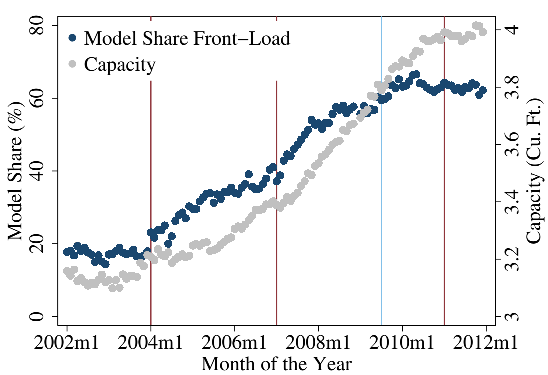
(c) Freezers

(d) Compact Refrigerators

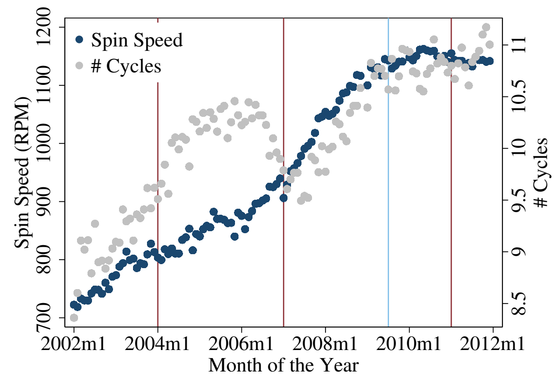
(e) Full-Size Refrigerators

Figure 4. Price-Energy Efficiency-Adjusted Quality Indexes

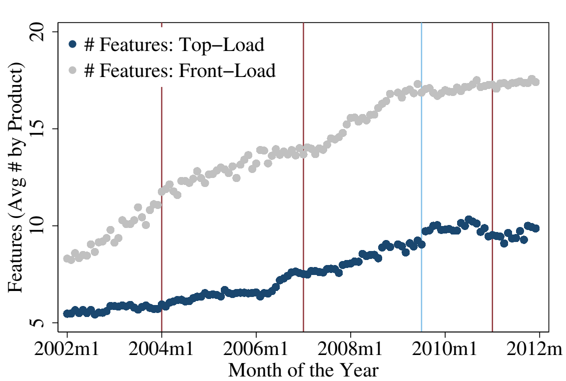
.



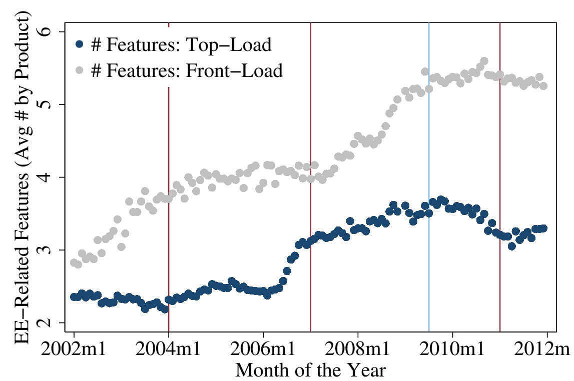
(a) Style and Size



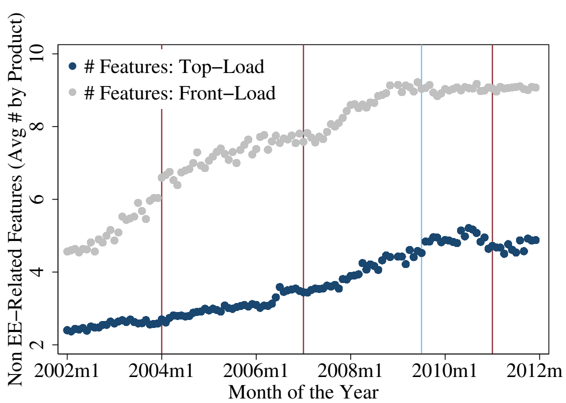
(b) Motor Speed and Washing Cycles



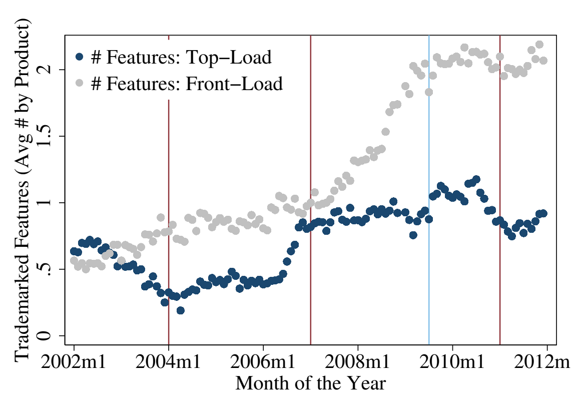
(c) All Features



(d) EE-Related Features



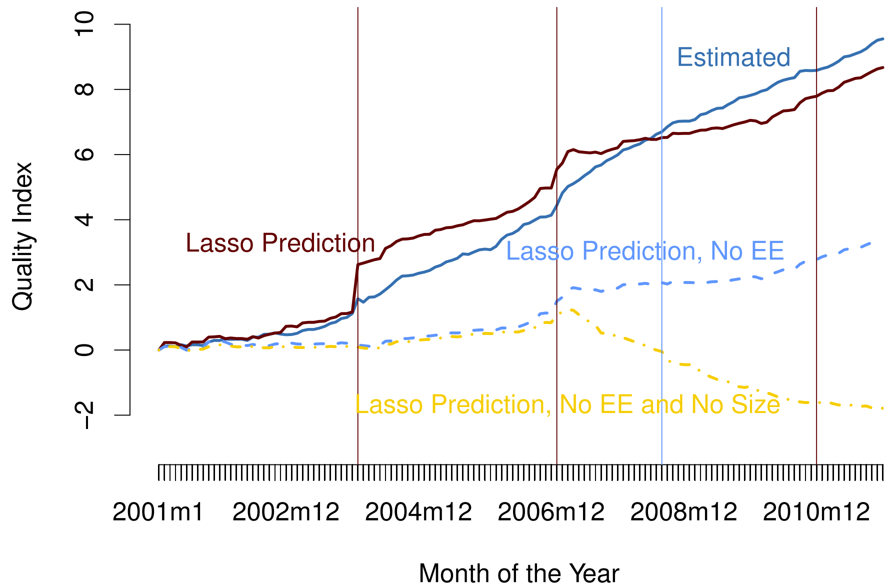
(e) Non EE-Related Features



(f) Trademarked Features

FIGURE 5. Evolution of Attributes and Features: Clothes Washers

.



(a) Clothes Washers

FIGURE 6. Energy Efficiency-Adjusted Quality Indexes

.

Table 1. Change in Price-Adjusted Quality

Year	Clothes Washers	Dishwashers	Room Air Conditioners	Refrigerators-Freezers
1988	(1) 1st MES	(1) 1st MES		
⋮				
1990			(1) 1st MES	(1) 1st MES
⋮				
1993				(1) 2nd MES
1994	(1) 2nd MES	(1) 2nd MES		
⋮				
1996		(6) ES start	(10) ES start	(6) ES start*
1997	(5) ES start			
⋮				
2000			(10) ES; (10) 2nd MES	
2001	(1) ES	(1) ES		(1) ES*; (7) 3rd MES
⋮				
2003		(9) TP	(10) ES⋄	(1) ES†
2004	(1) 3rd MES Tier 1; (1) ES			(1) ES*
2005			(11) ES⋄⋄	
⋮				
2007	(1) 3rd MES Tier 2; (1) ES	(1) ES		
2008				(4) ES*
2009	(7) ES	(8) ES		
2010		(1) 3rd MES		
2011	(1) 4th MES; (1) ES			
2012	(4) TP	(1) ES; (12) TP		
2013	(2) ES	(1) 4th MES	(10) ES	
2014		(1) ES	(6) 3rd MES; (5) ES	(9) ES; (4) TP; (9) 4th MES

Notes: The Table describes the timing of past federal regulatory action taken for clothes washers (CW), room air conditioners (RACs), refrigerators (REF) and dishwashers (DW). Legend: Numbers (1st, 2nd, etc.) = the order of the federal minimum efficiency standards (MES) effective dates; ES = ENERGY STAR criteria changed; and TP = Test Procedure changed; month effective shown in parentheses. Additionally, for refrigerators the symbol * signifies changes to the ES policy that affected full-sized refrigerators only, while the symbol † signifies changes that affected compact refrigerator/freezers and freezers only. ⋄ The 2003 RAC standard was only an expansion to cover RACs without louvered sides. ⋄⋄ The 2005 RAC standard was only an expansion to cover RACs with reverse cycle.

Table 2. Change in Price-Adjusted Quality

Dep. Variable: Price-Adjusted Quality Index		First-Diff		Diff-in-Diff		First-Diff w. Matching		Diff-in-Diff w. Matching	
Revision Event	Non-ES vs. ES	ΔQI	t-stat	ΔQI	t-stat	ΔQI	t-stat	ΔQI	t-stat
CW: MES 01/2004	Non-ES	1.30	3.94	0.96	3.16	0.89	2.61	0.47	1.52
	ES	3.03	8.46	2.68	8.13	2.02	4.36	1.58	3.81
CW: MES 01/2007	Non-ES	3.29	9.95	3.05	10.29	3.09	8.15	3.16	9.52
	ES	3.31	8.39	3.25	9.21	2.08	4.35	2.19	5.23
DW: MES 01/2010	Non-ES	1.17	2.79	0.49	1.30	1.05	2.34	0.41	1.04
	ES	-1.34	-1.76	-1.93	-2.85	0.09	0.11	-0.90	-1.30
CW: ES 01/2009	ES	5.03	5.02	4.63	5.21	4.89	3.93	4.70	4.32
DW: ES 01/2007	ES	0.52	1.46	0.43	1.32	0.40	1.06	-0.04	-0.12
DW: ES 01/2009	ES	1.94	5.77	1.34	4.42	0.78	2.09	1.30	3.97
REF: ES 01/2004	ES	0.45	0.89	0.18	0.39	0.19	0.33	-0.14	-0.29
REF: ES 04/2008	ES	2.01	9.45	1.73	9.07	2.17	9.94	2.00	10.41
Appliance Type FE		Yes		Yes		Yes		Yes	
Product Line FE		No		No		Yes		Yes	
Year FE		No		Yes		No		Yes	

Notes: Clothes Washers (CW); Dishwashers (DW), Full-Size Refrigerators (REF). The t-statistic that rejects the null hypothesis (the coefficient is not different than zero) at the 5% level takes a value of $+/-1.97$. All t-stats greater than $|1.97|$ identifies coefficients significant at the 5% level.

TABLE 3. Changes in Prices and Electricity

Dep. Variable: log(price)		First-Diff		Diff-in-Diff		First-Diff w. Matching		Diff-in-Diff w. Matching	
Revision Event	Non-ES vs. ES	ΔQI	t-stat	ΔQI	t-stat	ΔQI	t-stat	ΔQI	t-stat
CW: MES 01/2004	Non-ES	-0.14	-2.17	-0.08	-1.17	-0.09	-1.31	-0.03	-0.38
	ES	0.30	4.43	0.36	5.21	-0.04	-0.49	0.02	0.21
CW: MES 01/2007	Non-ES	0.12	1.85	0.11	1.66	0.09	1.27	0.07	0.98
	ES	0.04	0.54	0.03	0.34	-0.04	-0.44	-0.07	-0.66
DW: MES 01/2010	Non-ES	0.08	0.97	0.04	0.44	0.09	1.04	0.06	0.65
	ES	-0.64	-4.50	-0.71	-4.93	-0.04	-0.27	-0.09	-0.62
CW: ES 01/2009	ES	0.16	0.86	0.17	0.88	-0.01	-0.03	0.00	0.00
DW: ES 01/2007	ES	0.08	1.15	0.07	1.00	-0.03	-0.41	-0.06	-0.78
DW: ES 01/2009	ES	0.06	0.88	0.09	1.35	0.00	0.00	0.03	0.46
REF: ES 01/2004	ES	0.07	0.69	0.11	1.14	0.07	0.74	0.11	1.22
REF: ES 04/2008	ES	0.10	2.52	0.08	1.98	0.10	2.72	0.08	2.15
Appliance Type FE		Yes		Yes		Yes		Yes	
Product Line FE		No		No		Yes		Yes	
Year FE		No		Yes		No		Yes	
Dep. Variable: log(kWh/y)									
CW: MES 01/2004	Non-ES	-0.52	-15.13	-0.52	-15.15	-0.54	-14.97	-0.54	-15.17
	ES	-0.46	-8.29	-0.46	-8.49	-0.42	-6.13	-0.42	-6.38
CW: MES 01/2007	Non-ES	-0.35	-8.98	-0.35	-9.24	-0.39	-8.72	-0.39	-9.21
	ES	-0.45	-11.52	-0.45	-11.96	-0.20	-3.74	-0.21	-4.05
DW: MES 01/2010	Non-ES	-0.20	-4.18	-0.16	-3.53	-0.13	-1.96	-0.07	-1.14
	ES	-0.02	-0.31	0.03	0.48	-0.04	-0.52	0.04	0.47
CW: ES 01/2009	ES	-0.54	-5.53	-0.52	-5.64	-0.29	-2.25	-0.28	-2.26
DW: ES 01/2007	ES	-0.16	-4.44	-0.11	-3.21	-0.15	-3.80	-0.12	-3.10
DW: ES 01/2009	ES	-0.05	-1.50	-0.08	-2.50	-0.06	-1.50	-0.10	-2.54
REF: ES 01/2004	ES	-0.04	-0.79	-0.04	-0.76	-0.04	-0.81	-0.04	-0.79
REF: ES 04/2008	ES	-0.10	-4.86	-0.11	-5.33	-0.10	-5.01	-0.11	-5.47
Appliance Type FE		Yes		Yes		Yes		Yes	
Product Line FE		No		No		Yes		Yes	
Year FE		No		Yes		No		Yes	

Notes: Clothes Washers (CW); Dishwashers (DW), Full-Size Refrigerators (REF). The t-statistic that rejects the null hypothesis (the coefficient is not different than zero) at the 5% level takes a value of $+/-1.97$. All t-stats greater than $|1.97|$ identifies coefficients significant at the 5% level.

TABLE 4. Change in Price-EE-Adjusted Quality

Dep. Variable: **EE-Price-Adjusted Quality Index**		First-Diff		Diff-in-Diff		First-Diff w. Matching		Diff-in-Diff w. Matching	
Revision Event	Non-ES vs. ES	ΔQI	t-stat	ΔQI	t-stat	ΔQI	t-stat	ΔQI	t-stat
CW: MES 01/2004	Non-ES	-0.52	-1.58	-0.67	-2.00	-1.03	-3.37	-1.05	-3.39
	ES	0.99	1.91	0.88	1.72	0.25	0.45	0.23	0.42
CW: MES 01/2007	Non-ES	3.16	8.83	2.98	8.54	2.59	6.92	2.66	7.41
	ES	2.80	7.76	2.72	7.72	2.56	6.07	2.64	6.52
DW: MES 01/2010	Non-ES	0.92	2.17	0.57	1.37	0.26	0.50	-0.09	-0.19
	ES	-1.18	-1.70	-1.62	-2.41	0.01	0.02	-0.69	-1.04
CW: ES 01/2009	ES	4.59	5.02	4.40	4.98	4.51	4.14	4.44	4.27
DW: ES 01/2007	ES	0.36	1.09	0.32	0.97	0.13	0.38	0.08	0.24
DW: ES 01/2009	ES	1.31	4.22	1.21	3.97	0.63	1.94	1.04	3.30
REF: ES 01/2004	ES	0.34	0.72	0.22	0.48	-0.43	-0.85	-0.44	-0.91
REF: ES 04/2008	ES	1.76	9.01	1.52	7.98	1.88	9.81	1.75	9.39
Appliance Type FE		Yes		Yes		Yes		Yes	
Product Line FE		No		No		Yes		Yes	
Year FE		No		Yes		No		Yes	

Notes: Clothes Washers (CW); Dishwashers (DW), Full-Size Refrigerators (REF). The t-statistic that rejects the null hypothesis (the coefficient is not different than zero) at the 5% level takes a value of $+/-1.97$. All t-stats greater than $|1.97|$ identifies coefficients significant at the 5% level.

TABLE 5. OLS and LASSO Regressions

	OLS Top-Load		OLS Front-Load		LASSO Top-Load		LASSO Front-Load	
Feature	Est.	s.e.	Est.	s.e.	Est.	s.e.	Est.	s.e.
kWh/y	-0.01	-9.16	-0.02	-3.47	-0.01	-10.51	-0.01	-4.07
Size (Cu. Ft.)	1.29	3.02	2.30	4.49	1.35	3.68	2.37	5.50
# Cycles	-0.08	-2.32	0.02	0.58	-0.08	-2.61	0.01	0.32
Add a Garnment	1.79	0.58	0.16	0.20	0.64	0.39		
Advanced Motor Features	8.72	1.42	2.86	2.09	0.41	0.27	2.35	3.26
Automatic Timer			0.91	0.96				
Bleach Dispenser	-1.59	-1.09	-0.09	-0.05	-1.62	-1.46		
Clean Action	0.02	0.03	-0.55	-1.00				
Cold Temperature Default	-0.54	-0.56	1.20	1.61			1.21	2.23
Cycle Status Remaining Time	2.72	1.66	-3.36	-3.99			-2.68	-4.12
Cycle Status End Signal	-0.38	-0.82	2.47	3.27	-0.22	-0.55	1.57	3.09
Cycle Status Lights	2.13	3.04	1.11	1.94	2.28	3.92	0.53	1.29
Delay Start	2.32	1.36	-0.24	-0.37			-0.11	-0.19
Detergent Dispenser	-2.05	-1.11	-0.86	-0.40			-0.37	-0.28
Dryer Ready			-1.28	-1.03			-1.22	-1.04
ES-certified	-0.59	-1.01	-0.49	-1.13				
Electronic Control	-2.76	-2.91	-1.76	-1.74	-2.07	-3.19	-1.62	-2.06
Extra Rinse	-0.19	-0.43	-0.35	-0.56	-0.21	-0.51		
Fabric Softener Dispenser	-0.48	-0.55	-0.20	-0.14	-0.54	-0.70	-0.14	-0.22
Heater	-1.43	-0.44	-0.01	-0.02				
Injection Dispenser	1.04	0.54	-0.30	-0.49			-0.23	-0.49
Intercept	11.17	3.83	7.18	1.51				
Maximum Spin Speed	-1.09	-0.67	-0.20	-0.19			-0.29	-0.37
NSF Certified	1.65	0.89	-0.37	-0.62	1.20			
Other Dispenser			-1.12	-1.19			0.14	0.24
Other Features Tub			0.18	0.20			0.29	0.36
Programmable Control	0.36	0.24	0.76	1.74			0.68	1.77
Quickwash	-0.29	-0.24	0.39	0.94	-0.91	-1.02	0.52	1.41
Remote Laundry Monitoring			2.61	1.79			2.29	2.15
Sanitize Cycle	1.63	1.21	1.26	2.78	1.74	1.74	1.13	2.99

Notes: The t-statistic that rejects the null hypothesis (the coefficient is not different than zero) at the 5% level takes a value of $+/-1.97$. All t-stats greater than $|1.97|$ identifies coefficients significant at the 5% level.

Table 6. OLS and LASSO Regressions, Contd.

	OLS Top-Load		OLS Front-Load		LASSO Top-Load		LASSO Front-Load	
Feature	Est.	s.e.	Est.	s.e.	Est.	s.e.	Est.	s.e.
Sanitize Heat	-13.72	-1.41	0.18	0.23				
Sanitize Silver Ion			0.64	0.53			0.76	1.02
Sanitize Steam Technology			0.78	1.39			0.41	0.97
Spin Time Option	0.34	0.14			0.10	0.07		
Smooth Balance	1.19	1.31	-0.09	-0.14	0.50	0.66	-0.36	-0.76
Smooth Noise	-9.14	-1.99	0.88	0.47	-1.27	-1.18		
Smooth Suspension	2.45	0.89	0.14	0.19			0.30	0.51
Soil Level Selector	2.23	2.47	4.05	5.45	1.64	2.17	3.32	5.95
Soil Level Sensor	2.96	2.71	0.95	0.70	2.96	4.48		
Special Door Access	-0.31	-0.10	0.29	0.25				
Spin Speed Option	-0.26	-0.48	1.18	1.48			0.98	1.47
Stainless Tub	3.23	0.87	-0.25	-0.39			0.00	-0.01
Temperature Selection	-2.47	-0.75	-0.23	-0.04	-2.77	-1.22	-3.94	-1.47
Temperature Sensor	0.76	1.55	-1.82	-2.00			-1.48	-2.54
Water Level Selector	-1.21	-0.75	1.53	0.73	-0.44	-0.74		
Water Level Sensor	-0.82	-0.47	-0.21	-0.11				
Water Saving Technology			3.37	1.53			0.92	0.63
Brand: Samsung			-9.73	-3.60			-5.98	-4.56
Brand: Whirlpool	0.61	0.83	-3.13	-1.09	0.28	0.52	0.61	1.21
Brand: Estate	0.59	0.63						
Brand: Bosh			-3.26	-1.07				
Brand: LG	-14.43	-1.35	-7.65	-3.66			-4.43	-4.34
Brand: GE	0.94	1.34	-5.61	-2.42	0.80	1.48	-2.89	-3.87
Brand: Frigidaire	1.03	1.35	-3.69	-1.70	0.73	1.18	-1.12	-1.71
Brand: Roper	-1.08	-1.24			-1.36	-1.83		
Brand: Maytag	1.23	1.60	-3.43	-1.25	0.97	1.74		

Notes: The t-statistic that rejects the null hypothesis (the coefficient is not different than zero) at the 5% level takes a value of $+/-1.97$. All t-stats greater than $|1.97|$ identifies coefficients significant at the 5% level.

Table 7. Change in Quality: LASSO Index

Dep. Variable: log(price)		First-Diff		First-Diff w. Matching	
Revision Event	Non-ES vs. ES	ΔQI	t-stat	ΔQI	t-stat
Price-Adjusted Quality Index: Predicted LASSO					
CW: MES 01/2004	Non-ES	3.36	8.18	3.45	8.30
	ES	2.04	3.12	2.55	3.32
CW: MES 01/2007	Non-ES	2.97	6.54	2.94	5.71
	ES	2.07	4.55	1.56	2.71
CW: ES 01/2009	ES	3.65	3.18	3.85	2.59
EE-Price-Adjusted Quality Index: Predicted LASSO					
CW: MES 01/2004	Non-ES	0.33	1.39	0.43	1.78
	ES	0.09	0.24	0.31	0.68
CW: MES 01/2007	Non-ES	1.65	6.27	1.74	5.76
	ES	1.21	4.60	1.21	3.59
CW: ES 01/2009	ES	1.77	2.64	2.17	2.49
Size-EE-Price-Adjusted Quality Index: Predicted LASSO					
CW: MES 01/2004	Non-ES	0.30	1.30	0.29	1.26
	ES	0.19	0.52	0.21	0.49
CW: MES 01/2007	Non-ES	-0.19	-0.76	-0.10	-0.34
	ES	-1.13	-4.41	-1.28	-3.99
CW: ES 01/2009	ES	-1.33	-2.06	1.26	1.52

Appendix A. Additional Details: Policy Background Appliance Market

This appendix provides detailed information about the minimum efficiency standards and ES standards for the different appliance categories covered in our empirical analysis. Although we distinguish explicitly among six different categories, in the regulatory process some of these categories are grouped together and are considered different product classes that belong to a broader category. For instance, this is the case for compact refrigerators, freezers, and full-size refrigerators.

Clothes Washers: The standards for clothes washers effective in 1994 were based on the energy factor (EF), which was a measure of cubic feet of capacity per kwh per cycle, and included energy consumed by the washer directly, as well as the energy required to heat the water. The standard effective in 1994 only applied to top-load washing machines. The next standards, which were adopted in 2001, became effective in a two tier process in 2004 and 2007. These standards were based on the modified energy factor (MEF), which expanded upon the EF by also incorporating a measure of the energy necessary to dry the clothes at the completion of the wash cycle. This captured the fact that washers that were more effective at spinning moisture out of the clothing were more energy efficient because the clothing required less energy to be dried thereafter. The standards that rolled out in 2004 and 2007 were applied to both top-load and front-load washers, and while these different configurations were recognized in the rulemaking analysis, the level of the standard established based on the MEF did not differ between these two product types. The clothes washer standard effective in 2011 was based on the MEF as described above, but was now also based on the water factor (WF), which captured the water use of the washer. The standard levels set for clothes washers vary based on the size of the washer: compact (less than 1.6 cubic feet) versus standard size washers. There have also been elements of technology-based standards for clothes washers. In particular, NAECA required that, effective in 1988, clothes washers must be manufactured with an unheated rinse option.

Dishwashers: The standard for dishwashers effective in 1994 was based on the energy factor (EF) for dishwashers, which was a measure of cycles per kWh. The second NAECA standard for dishwashers was delayed by the Process Rule.[35] Once the next revisions of the dishwasher standard

[35] The "Process Rule" refers to the 1995-96 Department of Energy (DOE) review of the process for developing appliance standards, which resulted in the suspension of several rulemakings ("Procedures for Consideration of New or Revised Energy Conservation Standards for Consumer Products," 61 FR 36974 (July 15, 1996) 10 CFR 430 Appendix A to Subpart C, or the "Process Rule").

came around (prompted by EISA in 2007), the metric changed for the standard effective in 2010 to that of kWh per year. The standard effective for dishwashers in 2013 was the result of a direct final rule as part of the Joint Petition agreement.[36] The product classes by which standard levels are differentially set for dishwashers are also based on size: compact versus standard sized dishwashers.

Room Air Conditioners: The metric by which the standards have been set for room air conditioners is the energy efficiency ratio (EER), which is a measure of cooling capacity (in Btus per hour) by electric input power (in watts). The second room air conditioner standard was similarly delayed by the Process Rule (see dishwashers above). The product classes for room air conditioners are based on capacity (in Btus per hour); whether or not the air conditioner has louvered sides; whether it is "reverse cycle," or has a heating cycle; and whether it is intended for window versus wall installation. The standard for room air conditioners effective in 2014 was also a part of the Joint Petition agreement (see dishwashers, above).

Refrigerators/Freezers: The standards for refrigerators and freezers effective in 1990 were not directly linked to the size of the unit, the standards thereafter were set as a function of the adjusted volume (AV), in cubic feet, of the unit. The metric the standards are based on is kWh per year, but the level of the standard cutoff for different units varies based on the AV. Refrigerators and freezers have by far the most separate product classes of all the appliances covered by standards. They are differentiated based on size (compact versus standard); freezers with manual versus automatic defrost; the mounting orientation of combined refrigerator-freezers (top-mount, side-by-side, or bottom-mount freezers); standalone freezers that are upright versus chest; and top-mount and side-by-side with and without through the door (TTD) ice dispensing. For example, all told the standard effective in 2001 had 18 separate product classes for refrigerators and freezers, with two more added in 2005 following a petition by stakeholders. The standard levels for all of these product classes then varied by AV.

Appendix B. **Clustering Analysis and Nested Logit Estimation**

For the nested logit estimation, we determine the nest structure for each appliance category, by performing a clustering analysis on a subset of observed product characteristics. The algorithm

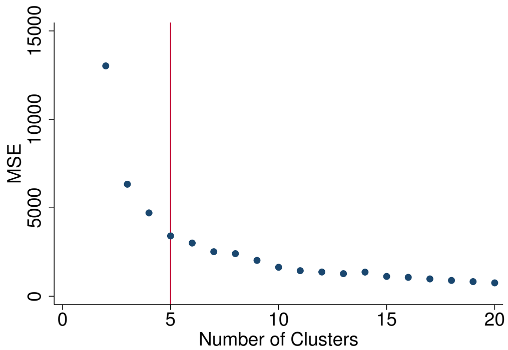
[36]"Agreement on Minimum Federal Efficiency Standards, Smart Appliances, Federal Incentives and Related Matters for Specified Appliances" (the "Joint Petition" or "Consensus Agreement").

that we use is a k-means clustering analysis (Hastie, Tibshirani, and Friedman 2009) implemented in Stata with the command "cluster kmeans." A k-means clustering analysis consists of partitioning J observations, each described by a vector x of size K, which in our case consists of J products with K attributes, into G sets $\mathbf{S} = \{S_1, S_2, ..., S_G\}$ by minimizing a within-cluster mean square error (MSE):

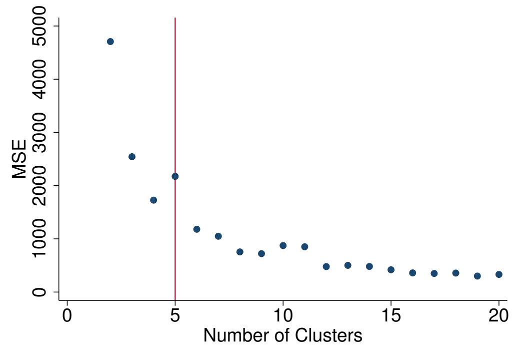
$$\arg\min_{\mathbf{S}} \sum_{g=1}^{G} \sum_{\mathbf{x} \in S_g} \|\mathbf{x} - \overline{\boldsymbol{x}}_{\boldsymbol{g}}\|^2$$

where $\overline{x}_g$ represents the mean of the observations portioned in the set S_g. The algorithm takes as input the number of sets G and the K dimensions of the vector describing each observation. For each appliance category, we select the number of clusters, G, using the so-called "hockey stick" criteria. We computed 8 iteratively for different values of G and plotted the objective function as a function of G. As shown in Figure B.1, the function 8 drops sharply for low values of G and stabilizes at a kick point, which mimics the shape of a "hockey stick." The heuristic is to use G that corresponds to the kick point. Across all appliance categories, a value of $G = 5$ tends to be the most appropriate.

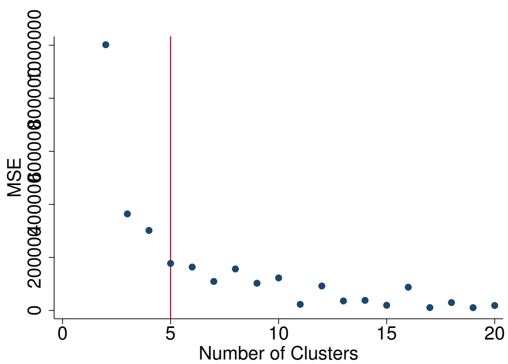
The K product attributes that we use for each appliance category were selected using a qualitative process. We identified the most prominent attributes and focused on attributes related to energy efficiency. Table B.1 lists the various attributes used for each appliance category.



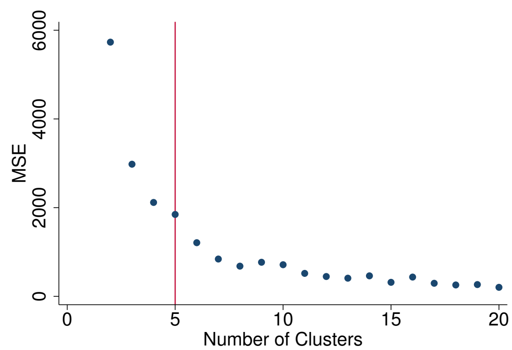
(a) Clothes Washers



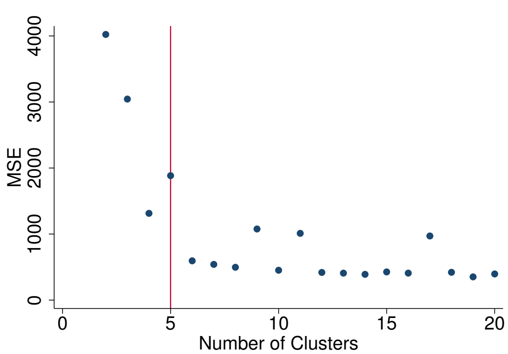
(b) Dishwashers



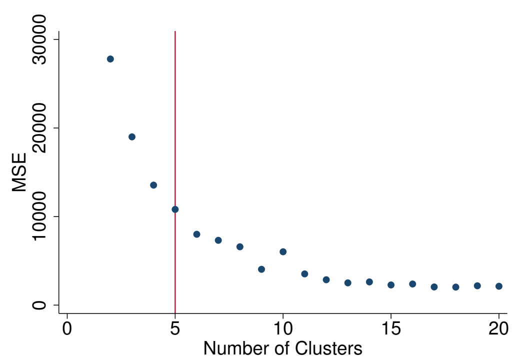
(c) Room ACs



(d) Freezers



(e) Compact Refrigerators



(f) Full-Size Refrigerators

Figure B.1. Clustering Analysis: MSE versus # of Clusters

TABLE B.1. Product Attributes for K-means Clustering Analysis

	Attributes Selected for K-means Clustering Analysis
Clothes Washers	price, capacity, electricity consumption, ES-certification, front-load versus top-load
Dishwashers	price, capacity, electricity consumption, ES-certification, built-in versus other
Room Air Conditionners	price, capacity, energy efficiency ratio, ES-certification, built-in versus other
Freezers	price, capacity, electricity consumption, ES-certification, upright versus chest freezer
Compact Refrigerators	price, capacity, electricity consumption, ES-certification, top-freezer versus other
Full-Size Refrigerators	price, capacity, electricity consumption, ES-certification, top-freezer, bottom-freezer or side-by-side

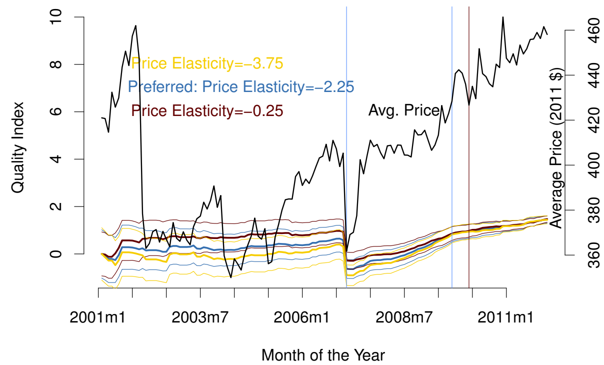
Appendix C. Sensitivity Tests

This Appendix shows the price-adjusted quality index estimated using different specifications and modeling assumptions. Overall, the results are similar to clothes washers presented in the main text. Controlling for product age tends to have the greater impact on the quality indexes. Note that when we do not control for product age, the quality index for some appliance category (freezers, compact and full-size refrigerators) displays large discrete jumps that are not coordinated with a standard revision event (Figure 2). These jumps are an artifact of the data collection technique of the NPD Group. From year to year, NPD may change the set of models that it is tracking and the retailers it is working with. In the present case, the fact that the discrete jumps occurred exactly at the beginning of a calendar year where we observe a large number of new models entering the sample suggests a change in the collection of the panel data. Qualitatively, controlling for product age is particularly important for dishwashers. Without product age, the 2007 revision in the ES standard suggests a large increase in quality, but we observe a large decrease once we account for product age.

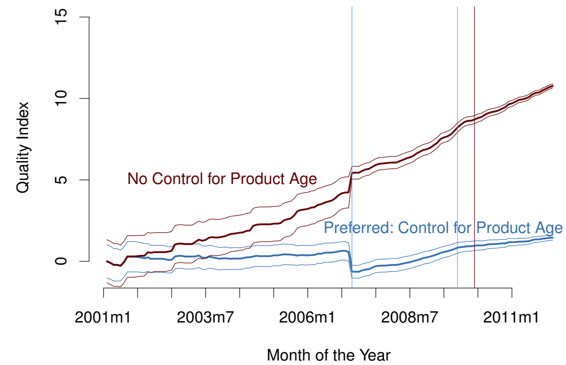
Similarly to clothes washers, for all five appliance categories, the effect of calibrating the coefficient on price is rather inconsequential. Only for full-size refrigerators, we observe a large and statistically significant difference on the quality index for different values of the parameter η. Qualitatively, the conclusion is the same, though. On average, the price-adjusted quality of full-size refrigerators has increased during the 2001-2011 period.

The controls for unobserved horizontal product differentiation have little effect for all appliance categories, which also mirrors the results for clothes washers. The only exception is for dishwashers in the post-2007 period, where the indexes computed under the logit assumption are larger and statistically different than the index computed with a nested logit and distance function as a control. If we increase the number of nests from 5 to 28, we have a richer control for unobserved horizontal product differentiation and the index becomes smaller. For this appliance category, during that particular time period, the effect of horizontal product differentiation was thus important empirically.

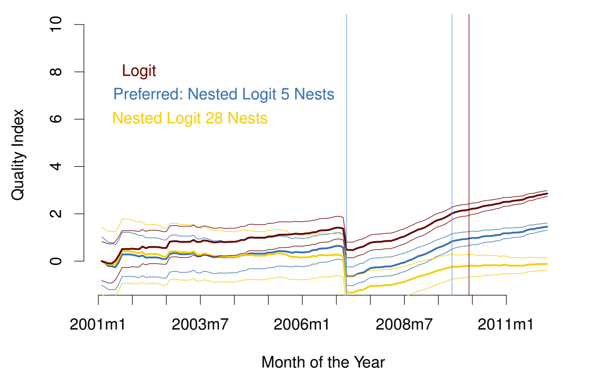
Finally, the sales-weighted quality indexes tend to display more month-to-month variability and display a slightly larger rate of increase relative to the model-weighted indexes. The overall patterns are, however, the same.



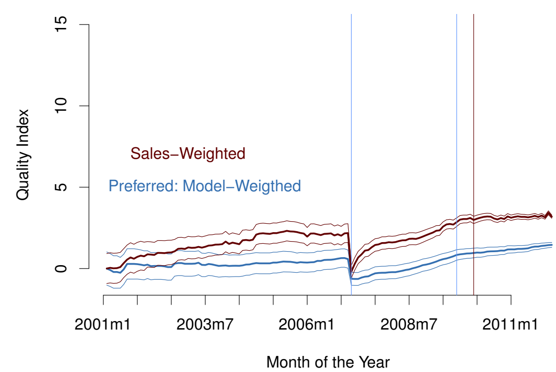
(a) Marginal Utility of Income



(b) Product Age



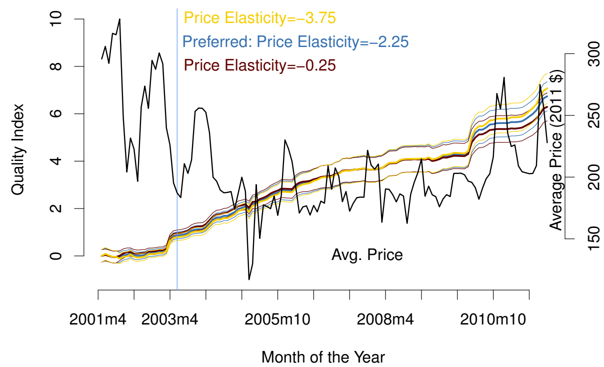
(c) Nested Logit Structure



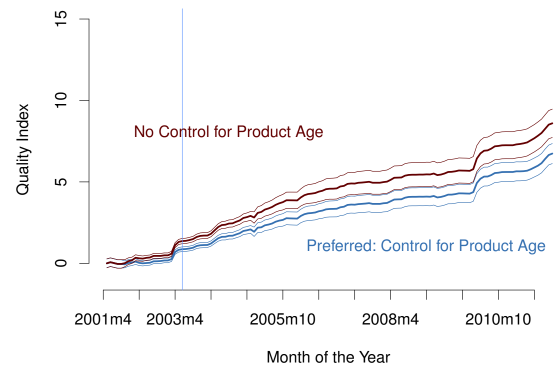
(d) Sales vs. Model Weighted

Figure C.1. Sensitivity Tests: Dishwashers

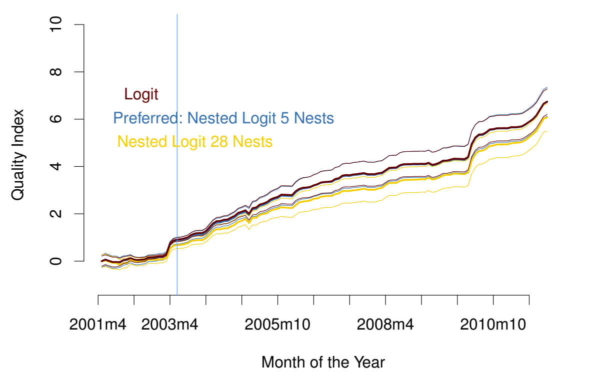
.



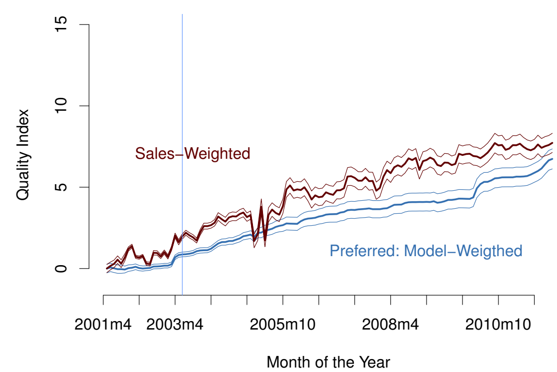
(a) Marginal Utility of Income



(b) Product Age



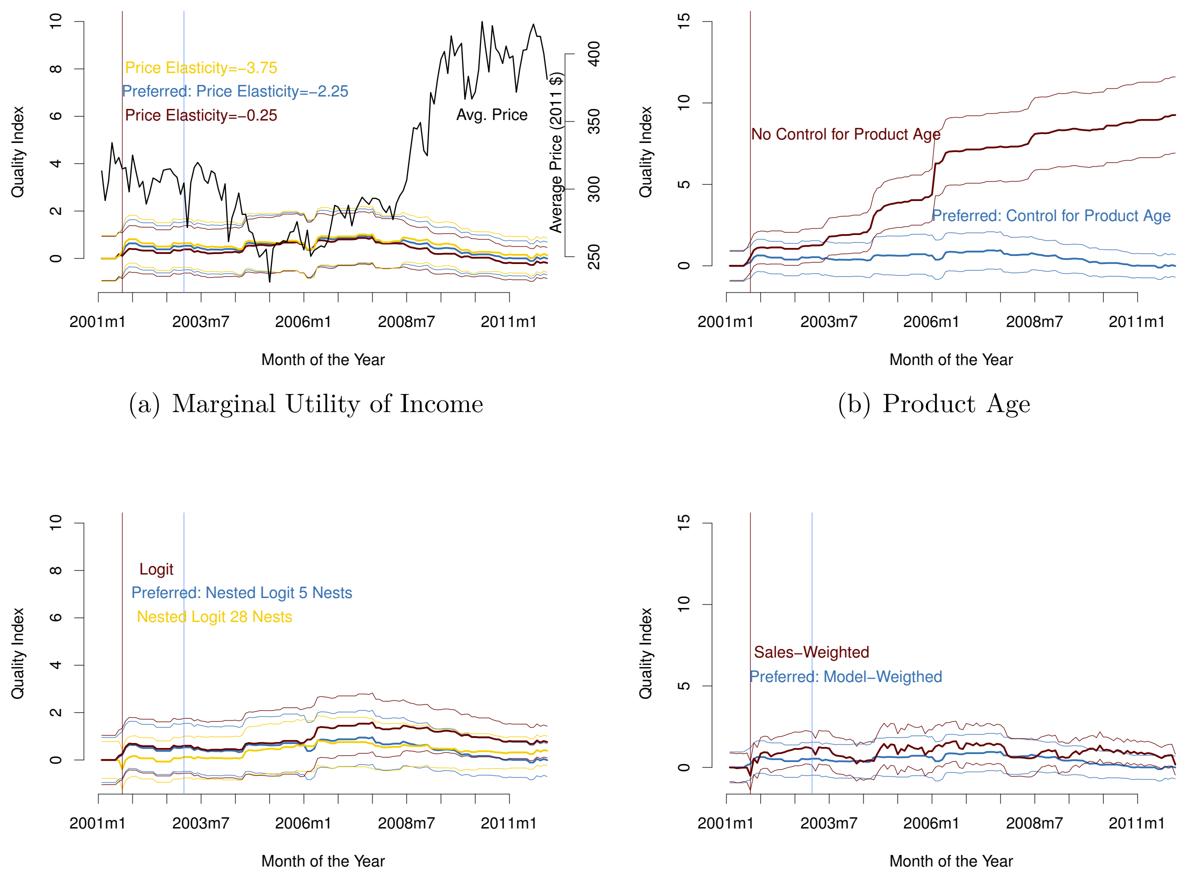
(c) Nested Logit Structure



(d) Sales vs. Model Weighted

Figure C.2. Sensitivity Tests: Room ACs

.



(a) Marginal Utility of Income

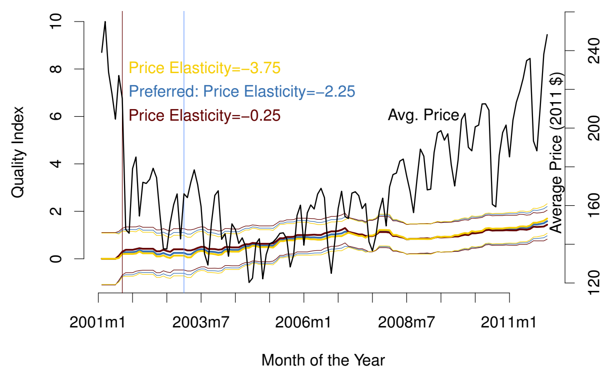
(b) Product Age

(c) Nested Logit Structure

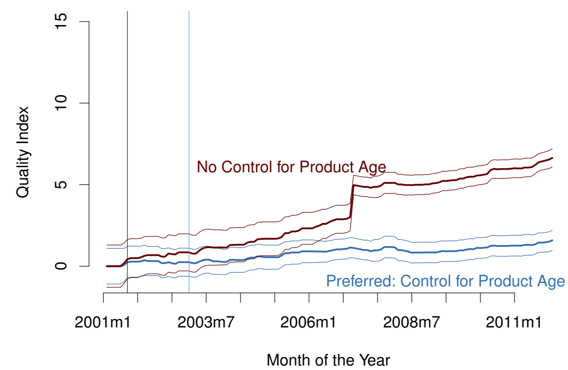
(d) Sales vs. Model Weighted

Figure C.3. Sensitivity Tests: Freezers

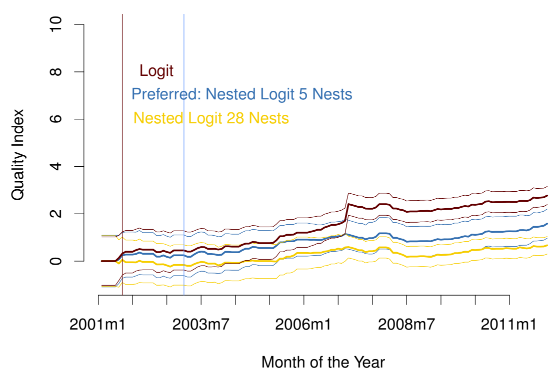
.



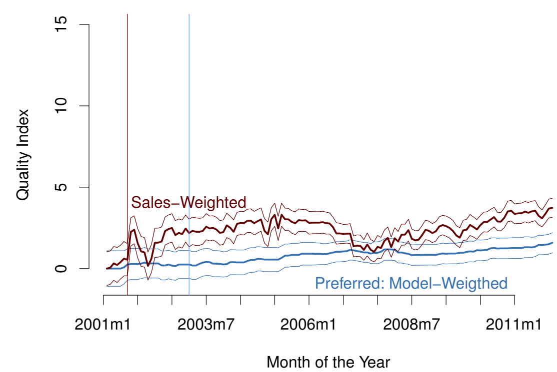
(a) Marginal Utility of Income



(b) Product Age



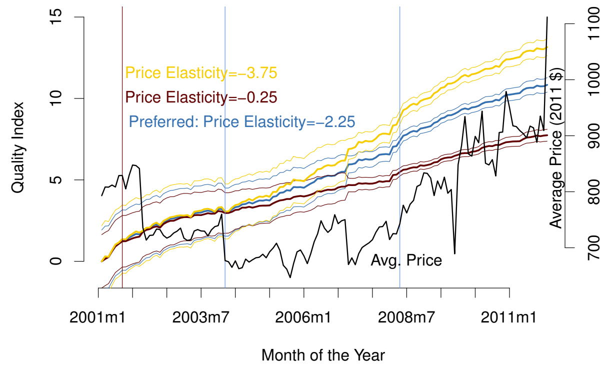
(c) Nested Logit Structure



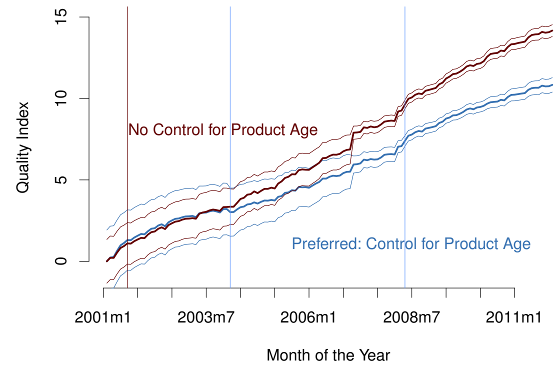
(d) Sales vs. Model Weighted

FIGURE C.4. Sensitivity Tests: Compact Refrigerators

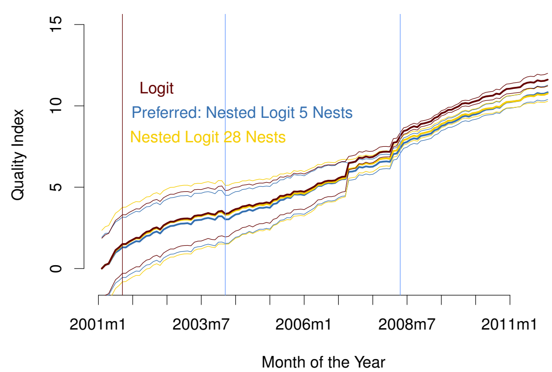
.



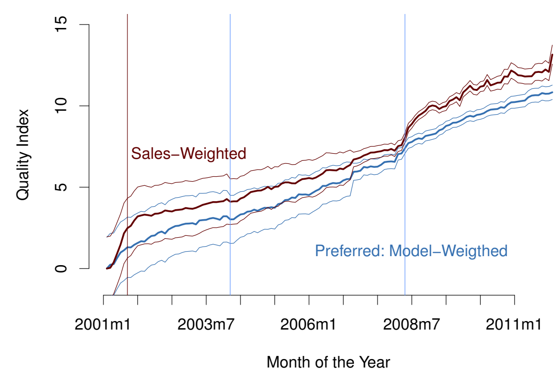
(a) Marginal Utility of Income



(b) Product Age



(c) Nested Logit Structure



(d) Sales vs. Model Weighted

FIGURE C.5. Sensitivity Tests: Full-Size Refrigerators

.

Table C.1. Change in Price-Adjusted Quality:
High Marginal Utility of Income $|\eta| = 0.0005$

Dep. Variable:		First-Diff		Diff-in-Diff		First-Diff		Diff-in-Diff	
Price-Adjusted Quality Index						w. Matching		w. Matching	
Revision Event	Non-ES vs. ES	ΔQI	t-stat	ΔQI	t-stat	ΔQI	t-stat	ΔQI	t-stat
CW: MES 01/2004	Non-ES	1.18	4.63	0.81	3.39	0.81	2.90	0.39	1.50
	ES	2.11	7.68	1.73	6.68	1.69	4.44	1.24	3.57
CW: MES 01/2007	Non-ES	2.65	10.44	2.58	11.04	2.37	7.64	2.50	8.96
	ES	2.47	8.16	2.51	9.03	1.63	4.15	1.79	5.07
DW: MES 01/2010	Non-ES	0.92	2.85	0.44	1.47	0.80	2.18	0.43	1.31
	ES	-0.11	-0.20	-0.53	-0.99	0.15	0.23	-0.46	-0.78
CW: ES 01/2009	ES	4.06	5.28	3.81	5.44	4.40	4.32	4.27	4.67
DW: ES 01/2007	ES	0.15	0.54	0.21	0.83	0.19	0.62	0.00	0.00
DW: ES 01/2009	ES	1.28	4.93	0.90	3.76	0.61	2.01	0.97	3.51
REF: ES 01/2004	ES	0.14	0.37	-0.15	-0.42	0.10	0.21	-0.24	-0.57
REF: ES 04/2008	ES	1.28	7.80	1.11	7.38	1.38	7.72	1.31	8.09
Appliance Type FE		Yes		Yes		Yes		Yes	
Product Line FE		No		No		Yes		Yes	
Year FE		No		Yes		No		Yes	

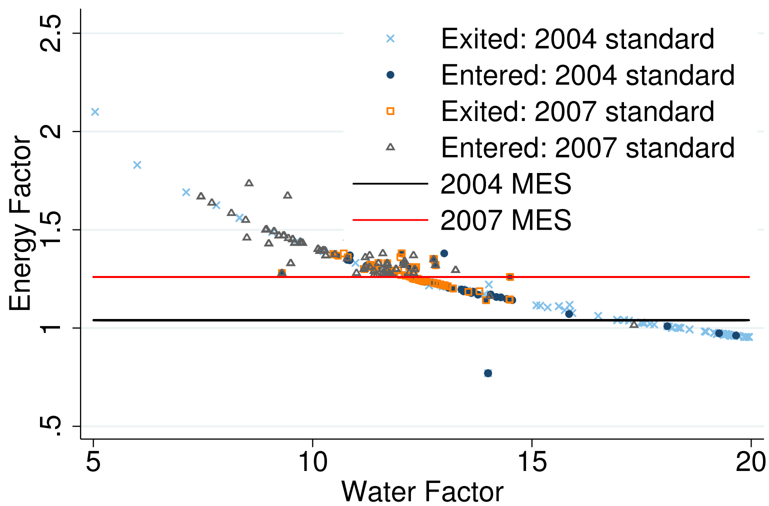
TABLE C.2. Change in Price-Adjusted Quality: Low Marginal Utility of Income $|\eta| = 0.0075$

Dep. Variable:		First-Diff		Diff-in-Diff		First-Diff		Diff-in-Diff	
Price-Adjusted Quality Index						w. Matching		w. Matching	
Revision Event	Non-ES vs. ES	ΔQI	t-stat	ΔQI	t-stat	ΔQI	t-stat	ΔQI	t-stat
CW: MES 01/2004	Non-ES	0.46	1.09	0.17	0.43	0.05	0.11	-0.34	-0.91
	ES	2.88	6.27	2.59	6.04	1.27	2.27	0.87	1.71
CW: MES 01/2007	Non-ES	4.24	9.98	3.86	10.03	4.17	9.09	4.18	10.28
	ES	4.32	8.53	4.17	9.10	2.84	4.88	2.88	5.59
DW: MES 01/2010	Non-ES	1.37	2.54	0.50	1.02	1.22	2.25	0.37	0.76
	ES	-2.29	-2.35	-3.05	-3.48	0.06	0.06	-1.24	-1.46
CW: ES 01/2009	ES	6.22	4.84	5.69	4.93	5.28	3.50	5.07	3.80
DW: ES 01/2007	ES	0.85	1.84	0.63	1.50	0.58	1.27	-0.05	-0.12
DW: ES 01/2009	ES	2.45	5.67	1.70	4.32	0.88	1.97	1.56	3.88
REF: ES 01/2004	ES	0.61	0.93	0.37	0.62	0.18	0.27	-0.13	-0.21
REF: ES 04/2008	ES	2.66	9.72	2.27	9.18	2.85	10.79	2.62	11.11
Appliance Type FE		Yes		Yes		Yes		Yes	
Product Line FE		No		No		Yes		Yes	
Year FE		No		Yes		No		Yes	

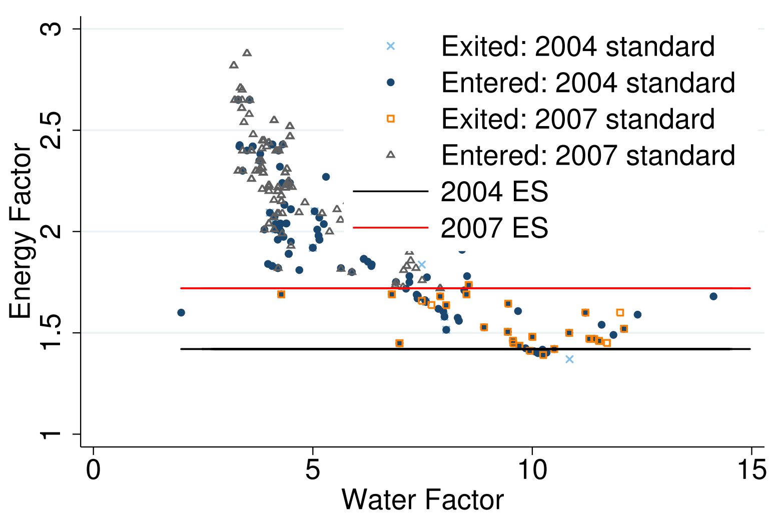
Table C.3. Interpretation of Changes in Price-Adjusted Quality: Money Metric

Revision Event	Non-ES vs. ES	First-Diff	Diff-in-Diff	First-Diff w. Matching	Diff-in-Diff w. Matching
Preferred: Marginal Utility Income $\|\eta\|$=0.0045, elasticity=-2.25					
CW: MES 01/2004	Non-ES	290	214	198	104
	ES	673	595	449	350
CW: MES 01/2007	Non-ES	732	677	686	703
	ES	736	721	463	488
DW: MES 01/2010	Non-ES	260	108	233	91
	ES	-297	-428	19	-201
CW: ES 01/2009	ES	1117	1028	1087	1045
DW: ES 01/2007	ES	116	95	89	-9
DW: ES 01/2009	ES	432	298	172	289
REF: ES 01/2004	ES	101	40	42	-32
REF: ES 04/2008	ES	447	384	482	445
Low Marginal Utility Income $\|\eta\|$=0.0075, elasticity=-3.75					
CW: MES 01/2004	Non-ES	62	23	6	-46
	ES	384	345	170	115
CW: MES 01/2007	Non-ES	565	515	556	558
	ES	576	556	378	384
DW: MES 01/2010	Non-ES	182	67	163	49
	ES	-305	-407	8	-165
CW: ES 01/2009	ES	830	759	704	676
DW: ES 01/2007	ES	113	84	78	-6
DW: ES 01/2009	ES	327	227	118	208
REF: ES 01/2004	ES	82	49	25	-17
REF: ES 04/2008	ES	354	303	381	349
High Marginal Utility Income $\|\eta\|$=0.0005, elasticity=-0.25					
CW: MES 01/2004	Non-ES	2353	1626	1622	776
	ES	4225	3468	3372	2483
CW: MES 01/2007	Non-ES	5304	5157	4738	5001
	ES	4949	5021	3255	3581
DW: MES 01/2010	Non-ES	1838	874	1595	864
	ES	-228	-1052	296	-913
CW: ES 01/2009	ES	8112	7621	8799	8534
DW: ES 01/2007	ES	297	422	388	-1
DW: ES 01/2009	ES	2556	1798	1223	1934
REF: ES 01/2004	ES	288	-301	191	-477
REF: ES 04/2008	ES	2555	2216	2758	2618

Appendix D. Entering and Exiting Models



(a) Non-ES Models



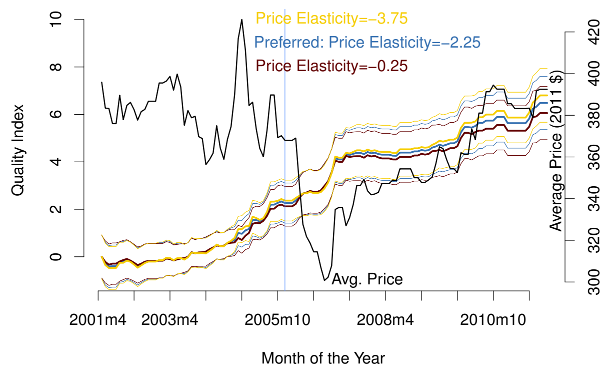
(b) ES Models

Figure D.1. Entering and Exiting Clothes Washer Models: 2004-2008

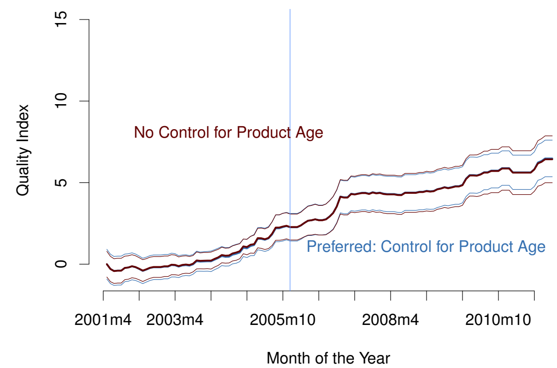
.

Appendix E. Additional Results: Room Air Conditioners with Reverse Cycle

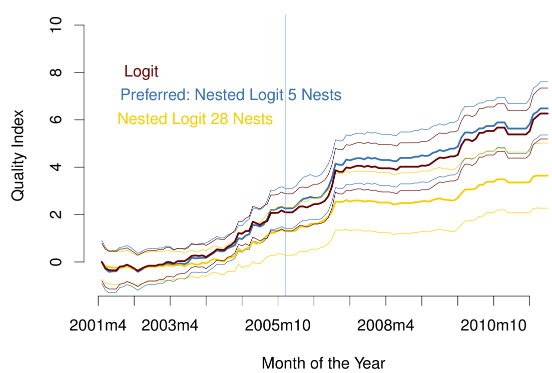
Room air conditioners with reverse cycle technology are commonly called heat pumps. They provide both heating and cooling. Prior 2005, this technology was covered by the ENERGY STAR certification. In our sample, we observe a relatively small number of room ACs with this technology. The analysis below was conducted on 119 models.



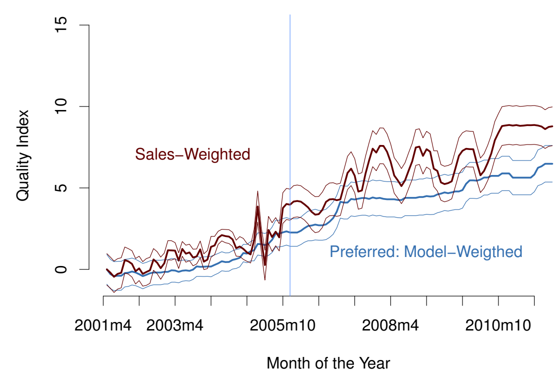
(a) Marginal Utility of Income



(b) Product Age



(c) Nested Logit Structure



(d) Sales vs. Model Weighted

Figure E.1. Sensitivity Tests: Room ACs with Reverse Cycle

.

Appendix F. LASSO Approach and Detailed Product Attributes: Clothes Washers

We implemented the LASSO estimation with the Post-LASSO algorithm. This is a two-stage estimation. First, we ran a LASSO regression using the price-adjusted quality indexes (estimated fixed effects) as the dependent variable. We used the whole vector of product attributes as regressors. We ran separate LASSO regressions for top-load and front-load models. We made this modeling decision after interacting the attributes with a dummy for front-load versus top-load. The model with the interaction terms shown that the dummy was not selected by the LASSO algorithm. The LASSO regressions selected a subset of attributes, others had a coefficient set to zero. The goal of the second step is to recover standard errors. The subset of product attributes selected by the LASSO regression are then regressed on the quality indexes with a simple OLS and the standard errors are computed.

TABLE F.1. Attributes of Clothes Washers

Attribute	Description	Coding	EE-Related	Subject to Trademark
kWh/y	Yearly Electricity Consumption Reported by Manufacturers to FTC	Continuous	Yes	
Size (Cu. Ft.)	Overall Capacity	Continuous		
# Cycles	Number of washing cycles	Continuous		
ES-certified	ENERGY STAR certified	0-1 Dummy		
NSF Certified	National Sanitation Foundation (NSF) certified	0-1 Dummy		
Brand	Brand dummies for Maytag, Roper, Samsung, Whirlpool, Bosch, Estate, Frigidaire, GE, and LG	0-1 Dummy		
Remote Laundry Monitoring		0-1 Dummy		
Add a Garnment		0-1 Dummy		Yes
Clean Action		0-1 Dummy		
Electronic Control		0-1 Dummy		
Programmable Control		0-1 Dummy		Yes
Cycle Status End Signal		0-1 Dummy		Yes
Cycle Status Remaining Time		0-1 Dummy		Yes
Cycle Status Lights		0-1 Dummy		Yes
Delay Start		0-1 Dummy		
Bleach Dispenser		0-1 Dummy		
Detergent Dispenser		0-1 Dummy		
Fabric Softener Dispenser		0-1 Dummy		
Injection Dispenser		0-1 Dummy	Yes	Yes
Other Dispenser		0-1 Dummy	Yes	
Special Door Access		0-1 Dummy		
Dryer Ready		0-1 Dummy		
Heater		0-1 Dummy	Yes	Yes
Water Level Selector		0-1 Dummy	Yes	Yes
Water Level Sensor		0-1 Dummy	Yes	
Water Saving Technology		0-1 Dummy	Yes	
Advanced Motor Features		0-1 Dummy	Yes	
Extra Rinse		0-1 Dummy		
Sanitize Heat		0-1 Dummy		Yes
Sanitize Silver Ion		0-1 Dummy		
Sanitize Steam Technology		0-1 Dummy		
Sanitize Cycle		0-1 Dummy		
Smooth Balance		0-1 Dummy		
Smooth Suspension		0-1 Dummy		
Smooth Suspension		0-1 Dummy		
Soil Level Selector		0-1 Dummy		Yes
Soil Level Sensor		0-1 Dummy	Yes	
Maximum Spin Speed		0-1 Dummy	Yes	Yes
Spin Speed Option		0-1 Dummy		
Spin Timer Option		0-1 Dummy		
Cold Temperature Default		0-1 Dummy	Yes	
Temperature Selection		0-1 Dummy	Yes	Yes
Temperature Sensor		0-1 Dummy	Yes	
Automatic Timer		0-1 Dummy	Yes	
Quickwash		0-1 Dummy	Yes	Yes
Other Features Tub		0-1 Dummy		
Stainless Tub		0-1 Dummy	Yes	