Beyond the Label: Do Ecodesign Regulations Reduce Real-World Energy Use?

Kai Gao*

Job Market Paper

November 2026

Click here for the latest version.

Abstract

This paper studies how ecodesign regulation shapes product design and consumer welfare when compliance is tied to a single operating mode rather than overall use. I examine the German washing machine market following the 2021 EU reform, which rescaled energy labels based on Eco-program efficiency and capped the program's duration to make Eco washes more convenient for consumers. Combining detailed product and wash-program data with an original household survey, I find that consumers prefer higher energy-efficiency ratings, yet use Eco for only about 10 percent of their wash cycles. Survey responses indicate that many interpret the label as measuring whole-machine efficiency. Meanwhile, firms offer more energy-efficient Eco programs but more energy-intensive Cotton programs. To explain these responses, I develop an equilibrium model of product entry and program design and estimate it using a novel approach combining composite conditional likelihood with Gibbs sampling, obtaining estimates comparable to those from full portfolio likelihood without enumerating each firm's large portfolio choice set. The estimates indicate that the duration cap shifts households that value cleaning performance toward Cotton programs, as shorter Eco cycles clean less thoroughly. This substitution increases firms' incentives to offer longer, more energy-intensive Cotton cycles. Counterfactual results indicate that, relative to a no-reform benchmark, the 2021 reform increases electricity consumption by 3.12 percent and reduces consumer welfare by € 231.1 million. Broadening the efficiency index to include Cotton programs can further increase consumer surplus and reduce electricity expenditure.

*I am grateful to Panle Jia Barwick, Jean-François Houde, Karam Kang, and Shangjun Li for their guidance and support. I thank Ashley Swanson for helpful comments, and Lorenzo Magnolfi, Hyuk-Soo Kwon, Alan Sorensen, Jonathan Kroah, and participants in the UW–Madison IO Lunch and IO Seminar for valuable feedback. I thank Gundula Czyzewski of BSH and Reiner Steininger of the University of Bonn for answering technical and industry questions, and Andrea Liu for assistance with collecting product-manual data. This project was supported by the Robock Scholarship at the University of Wisconsin–Madison. All errors are my own. E-mail: kgao33@wisc.edu.

1 Introduction

Many product regulations rely on standardized tests, even though their benefits depend on how products perform in actual use. Automobile emissions provide a prominent example: Volkswagen's use of defeat-device software allowed vehicles to satisfy laboratory tests while producing substantially higher emissions on the road (Barrett et al., 2015). A related test–use distinction arises in medicine, where drug efficacy in controlled clinical trials can differ from effectiveness in routine practice (Eichler et al., 2011; Rothwell, 2005). These examples raise a broader question: when regulation rewards outcomes measured under standardized conditions, how do responses along both measured and unmeasured dimensions shape its realized benefits?[1]

This paper studies mechanisms through which measured and realized performance can diverge, using the German washing-machine market as a setting. European energy-efficiency regulation evaluates washing machines using a standardized Eco program, while households choose among multiple programs in use, including Cotton and Easy-Care. Regulation can reduce Eco's energy consumption while also changing its washing performance, leading households to switch to unregulated programs. This substitution strengthens firms' incentives to improve washing performance in those programs through more energy-intensive designs, which can further encourage their use. The interaction between firms' program design and households' program choice can therefore offset energy savings in Eco and potentially increase realized energy use, even as measured efficiency improves.

I exploit a major reform of European washing-machine regulation implemented in 2021 (European Commission, 2019a). The reform changed both how energy efficiency was measured and the design constraints firms faced. First, the energy label was rescaled from the old classes A^{+++} through D to classes A through G, and the formula used to measure energy efficiency was revised. Because most products had accumulated in the highest classes under the old system, the rescaling restored differentiation in the label and increased the market reward to lowering measured Eco energy use. Second, the reform imposed a maximum duration on the Eco cycle to make the program more convenient to use. This constraint may, however, reduce washing quality and restrict an important energy-saving margin: lowering wash temperature and compensating for the resulting loss in washing quality with a longer cycle.[2] Together, these changes increase the return to improving measured Eco efficiency while restricting how firms can achieve those improvements, providing useful variation for studying firms' design responses.

To trace these responses across product design and household use, I assemble a novel dataset

[1] A related incentive problem arises in education, where test-based accountability can shift effort toward tested outcomes and away from dimensions of learning that are not measured (Jacob, 2005).

[2] European washing machines heat wash water electrically within the machine. Water heating is by far the largest component of energy use during a wash cycle, making lower wash temperatures an important source of energy savings.

that links product-market outcomes to program-level design and household behavior. The main data combine detailed model-level sales with hand-collected attributes from washing-machine manuals, including electricity use, water use, duration, and temperature for both regulated and unregulated wash programs. I complement these data with an original survey that I designed and commissioned in Germany, which measures households' program choices and their interpretation of the energy label. Together, these data allow me to study firms' design responses across programs, consumers' purchase and usage decisions, and the implications for realized energy consumption.

The data reveal three main findings. First, consumers respond strongly to the energy class but often misunderstand its scope. The 2021 label rescaling increases sales of models rated C or higher by 55 percent relative to those rated D or lower, while survey respondents often interpret the energy class as measuring the efficiency of the whole machine rather than the Eco program alone. Eco nevertheless accounts for only about 10 percent of reported wash cycles. Second, improvements in regulated Eco efficiency do not extend to unregulated programs. Within the same machine, Eco and Cotton energy use per cycle are negatively correlated, particularly for Cotton 60°C. Event-study estimates show that by the end of 2023, unweighted average energy use per cycle fell by 8.6 percent for Eco but rose by 14.9 percent for Cotton 60°C. Third, products more exposed to the Eco-duration cap exhibit larger changes in unregulated programs. Products whose pre-policy Eco durations exceeded the new cap by a larger proportion experience greater increases in Cotton energy use and duration relative to matched old-label-only products. A 10-percentage-point increase in this exposure is associated with 5.9 percent higher energy use and 2.1 percent longer cycles in Cotton 60°C.[3] These patterns are consistent with firms shifting design toward unregulated programs in response to constraints on Eco.

To explain these findings, I develop an equilibrium model of the washing machine market. Demand consists of two stages: product purchase and program use. At the purchase stage, consumers value the new energy class once the label becomes observable but do not directly value the energy efficiency of unregulated programs. They also anticipate the utility from the bundle of wash programs they will use. At the use stage, households choose among programs based on expected washing quality and convenience, which depend on program-specific inputs such as temperature, duration, and water use. The expected utility from program use enters product demand through a household-specific inclusive value, linking firms' program-design decisions to consumers' willingness to pay.

On the supply side, firms make discrete product-entry and continuous program-design decisions. Firms choose which products to introduce and their energy-efficiency classes, then select program-specific inputs to achieve washing-quality targets, subject to class-specific energy-efficiency require-

[3]Energy use also increases with exposure in Cotton 40°C and Easy-Care, although the estimated increases are smaller.

ments and duration constraints for Eco. Firms trade off the demand benefits of higher energy ratings and improved program quality against design costs and regulatory constraints.

I estimate demand in two steps. First, I estimate program-use preferences by simulated maximum likelihood, combining survey respondents' reported program-use frequencies with program inputs from product manuals matched to the washing machines they own. Variation in program inputs across machines and differences in household usage patterns identify preferences and their heterogeneity. Second, I use the estimated preferences to construct household-specific expected utility from each machine's bundle of wash programs and incorporate it into a random-coefficients model of product demand following Berry et al. (1995). The label rescaling provides additional variation in displayed energy classes, while the survey-based estimates allow me to account for consumers' valuation of program performance. Preferences for energy classes are identified by differences in sales across displayed classes, conditional on prices, program utility, other observed characteristics, and fixed effects. Additional variation in product characteristics and program designs comes from the entry of new models, which I assume firms choose before observing contemporaneous demand shocks.

On the supply side, I use policy-induced changes in program inputs to help identify program-design costs. The reform changes firms' incentives and constraints, and their design responses reveal the costs of adjusting program inputs. Conditional on observed product portfolios, I estimate these costs by minimum distance, matching moments describing input levels and changes at product introduction to their model counterparts, generated by firms' optimal design choices under engineering and regulatory constraints. Taking the estimated design costs as given, I then estimate entry costs in a static discrete-choice game, adopting an incomplete-information framework as in Seim (2006).

Estimating the entry game poses substantial computational challenges because each firm's portfolio choice set grows exponentially with the number of product lines. I develop an approach that combines composite conditional likelihood with Gibbs sampling to avoid enumerating this large choice set. The key assumption is that firms' private payoff shocks are independent and identically distributed type-I extreme value across feasible portfolios. This assumption yields exact conditional-logit probabilities in which the full portfolio normalizer cancels. Instead of evaluating full portfolio probabilities, I therefore use the probability of each line's entry choice—no entry or introduction in class A, B, C, or D—conditional on the firm's observed choices on its other lines. Expected profits still depend on rivals' full portfolio distributions, but a complete set of conditional entry probabilities uniquely determines each strictly positive joint distribution. I use these conditional probabilities to simulate rivals' portfolios through Gibbs sampling (Casella and George, 1992) and compute expected profits. I iterate between estimating entry parameters and updating beliefs, checking convergence in conditional choice probabilities. Monte Carlo exercises show that the approach recovers estimates comparable to those obtained using full portfolio likelihood. A further computational challenge is

that evaluating each candidate portfolio requires solving for optimal continuous program inputs. I address this by training a neural-network surrogate to predict optimal inputs and profits net of design costs, avoiding repeated constrained optimization during entry estimation.

The estimated model reveals substantial heterogeneity in how consumers trade off cleaning performance against cycle time. Relative to A^{+++} durations, a 30-minute reduction lowers Eco use by 5.5 p.p. among consumers who place greater value on cleaning performance but raises it by 0.3 p.p. among those who place greater value on saving time. On average, Eco's share falls by 1.65 p.p. and the combined Cotton share rises by 1.61 p.p. The product-demand estimates further show that consumers value both higher efficiency ratings and higher expected wash utility. These preferences give firms incentives to balance measured efficiency against the attractiveness of the full menu of wash programs. Under the same duration comparison, the average marginal operating profit from extending Cotton programs rises by about 19 percent, linking household substitution to firms' incentives to offer longer Cotton cycles.

Policy Implications I use the estimated model to evaluate three sets of policy changes: removing existing regulatory restrictions, broadening the efficiency index to include Cotton programs, and introducing an energy-based producer tax. The experiments examine how greater flexibility in Eco design and stronger incentives for efficiency outside Eco affect consumer welfare and electricity use. Firms adjust product introductions and program designs, while households adjust purchases and program use. I evaluate outcomes for purchase cohorts from 2020Q1 through 2023Q4, accounting for subsequent machine use by new purchasers and continued use of existing machines by nonbuyers. Unless otherwise stated, changes are measured relative to the model-implied current-policy equilibrium.

First, removing the Eco-duration cap while retaining energy labels and efficiency requirements allows firms to substitute washing time for temperature. Eco cycles become longer and cooler, while both Cotton programs become shorter and consume less electricity per cycle. These adjustments are consistent with longer Eco cycles weakening firms' incentives to serve households valuing cleaning performance through energy-intensive Cotton programs. The reported comparison shows a consumer-welfare gain of € 63.24 million and a 1.11 percent reduction in discounted electricity expenditure. Removing labels and efficiency-index restrictions while retaining the duration cap has a different effect: Eco becomes warmer and consumes more electricity per cycle, and savings in Cotton do not offset the increase in overall electricity expenditure. Consumer welfare rises by € 59.52 million, but discounted electricity expenditure increases by 0.51 percent. Together, these experiments distinguish the benefits of greater design flexibility from the consequences of weakening incentives for measured efficiency.

In a separate no-2021-reform benchmark, I hold the 2020Q1 product portfolio and program

designs fixed while allowing prices to adjust. Relative to this benchmark, the reported comparison shows that the reform raises discounted electricity expenditure by 3.12 percent and reduces consumer welfare by € 231.13 million. This comparison illustrates how improvements in tested efficiency can coexist with higher electricity costs once product design and household use are taken into account.

Second, broadening the efficiency index makes Cotton electricity consumption relevant to the displayed rating and strengthens firms' incentives to reduce it. I assign 80 percent of the weight to Eco and 10 percent to each Cotton program, while also removing the duration cap. This package reduces electricity consumption per cycle across all three programs, with the largest reduction in Cotton 40°C. Discounted electricity expenditure falls by € 21.26 million, and the reported consumer-welfare gain is € 76.59 million. Relative to removing the duration cap alone, incorporating Cotton into the index provides an additional consumer-welfare gain of € 13.35 million.

Third, an energy-based producer tax directly charges firms for the energy intensity of their products. The experiment replaces labels, efficiency-index restrictions, and the duration cap with a tax of € 50 per machine for each kWh of average per-cycle electricity consumption across Eco and the two Cotton programs. Although the tax encourages electricity savings, the simultaneous removal of labels and efficiency requirements weakens incentives for Eco efficiency. Eco electricity consumption per cycle rises, while consumption in the two Cotton programs falls. Together with changes in purchases and program use, these responses increase discounted electricity expenditure by 1.36 percent. Consumer welfare falls by € 74.75 million despite lump-sum rebates of tax revenue. Among the tax settings evaluated, zero tax maximizes consumer welfare. Under this specification, the tax package performs worse than removing the duration cap or broadening the efficiency index in terms of electricity expenditure and consumer welfare.

Related Literature This paper contributes to three strands of literature. First, it relates to research on multitask incentives and performance measurement. Holmstrom and Milgrom (1991) and Baker (1992) show that incentives based on incomplete performance measures can distort the allocation of effort across tasks. Empirical studies examine how performance-based incentives and quality disclosure affect the allocation of effort and investment across quality dimensions (Jacob, 2005; Lu, 2012; Gupta, 2021; Vatter, 2022). I extend this literature to environmental product regulation and identify a demand-mediated channel for multitasking distortions: consumers' responses to changes in regulated performance alter firms' returns to investment in unregulated dimensions. This channel can induce more energy-intensive designs even as measured efficiency improves.

Second, the paper contributes to research on the equilibrium effects of product regulation and endogenous attributes. Existing studies examine labeling and certification policies (Barahona et al., 2023; Houde, 2018, 2026; Davis and Metcalf, 2016) and environmental standards and attribute-based regulations (Reynaert, 2021; Reynaert and Sallee, 2021; Ito and Sallee, 2018; Barwick et al., 2024).

I examine how regulation targeting one operating mode affects both the design of other modes and consumers' allocation of use across them. Incorporating these two margins allows me to evaluate whether improvements in measured efficiency translate into lower actual energy consumption and higher consumer welfare.

Third, the paper contributes to empirical entry games and endogenous product variety. Seim (2006) estimates an incomplete-information entry game using equilibrium choice probabilities, while Jia (2008) exploits strategic structure to solve a game with interdependent entry decisions. Computation becomes challenging when many firms each choose among large portfolios of products. An alternative literature uses moment inequalities to estimate costs without solving the full portfolio game (Holmes, 2011; Houde et al., 2023; Eizenberg, 2014; Wollmann, 2018; Montag, 2023; Fan and Yang, 2025). These methods accommodate large action spaces but may deliver set identification, and counterfactuals require separate equilibrium computation. I develop a likelihood-based alternative under incomplete information. The key simplifying assumption is i.i.d. extreme-value private shocks at the portfolio level, which yields exact conditional-logit probabilities. This structure permits estimation and equilibrium simulation without full portfolio enumeration or restrictions such as supermodularity. The flexibility comes at the cost of assuming independent shocks across portfolios, including those sharing many products.[4] Under the maintained assumptions, the approach permits point estimation and uses the same equilibrium algorithm for estimation and counterfactuals. Its main computational advantage is handling many choices within each firm, with potential applications to product offerings, store locations, export destinations, sourcing, and transportation networks.

The rest of the paper proceeds as follows. Section 2 describes the institutional setting and the 2021 label reform, presents the data, and documents the main empirical patterns. Section 5 develops the equilibrium model. Section 6 describes the estimation strategy, and Section 7 presents the estimation results. Section 8 presents the counterfactual policy analysis. Section 9 concludes.

2 Institutional Background and Data

2.1 The 2021 EU Energy Label Reform

The EU energy-labeling framework covers a wide range of household appliances, including refrigerators, washing machines, dishwashers, and air conditioners. Its purpose is to provide consumers with standardized information on product performance, particularly energy efficiency, and to strengthen

[4]Portfolio-level shocks can capture private coordination costs or assessments of a product lineup. Their independence rules out shared unobserved product costs. With product-level shocks and interacting profits, conditioning on other chosen products generally changes the remaining shock's distribution, so the conditional-logit representation need not hold.

firms' incentives to improve product design. Retailers are required to display energy labels both in physical stores and in online listings. This paper focuses on household washing machines.

The EU has implemented three major energy-labeling regimes for household washing machines, introduced in 1995, 2011, and 2021. This paper studies the transition from the 2011 regime to the 2021 regime. The reform was announced in March 2019 and took effect on March 1, 2021. Transitional provisions effective in December 2019 allowed manufacturers to update control panels in anticipation of the new requirements, enabling firms to begin adapting their products in early 2020. Consumers, however, continued to see the old energy labels until March 2021, when the new labels were introduced. From March 2021, newly placed products had to comply with the revised ecodesign requirements, while existing inventory could remain on sale under transitional rules.[5]

The 2021 reform introduced two major changes. First, it rescaled energy-efficiency ratings and revised both the standardized testing procedure and the formula used to determine them. Second, it introduced additional restrictions on the design of the regulated program, particularly limits on cycle duration. Together, these changes altered both the information consumers observe when purchasing washing machines and the constraints firms face when designing wash programs. Figure A6 compares the labels under the two regimes.

Energy-Label Rescaling Washing machines offer multiple wash programs, including Cotton, Easy Care, and Quick Wash. The EU energy label, however, measures performance under standardized test cycles rather than across all available programs.[6] Throughout the paper, I refer to these regulated test cycles as the Eco program, whose definition differs across regimes. Under the 2011 regime, the Eco program refers to the Standard Cotton 60°C cycle at full and half load and the Standard Cotton 40°C cycle at half load.[7] Under the 2021 regime, the Eco program refers to Eco 40–60, designed to clean normally soiled cotton laundry washable at either 40°C or 60°C in a single cycle. Other wash programs, including regular Cotton, are available to consumers but do not enter the calculation of the energy-efficiency rating.

Energy-efficiency classes are determined by an Energy Efficiency Index (EEI), which compares energy consumption measured under the Eco program with a capacity-dependent reference level. By the late 2010s, most washing machines had clustered in the highest efficiency classes, reducing the label's ability to differentiate among products. To restore differentiation and allow room for future efficiency improvements, the 2021 reform revised the EEI formula and replaced the previous

[5]Retailers had 14 working days to replace existing labels. Certain discontinued models could continue to be sold with their old labels until November 30, 2021.

[6]Using designated test cycles facilitates comparability across products and limits certification costs, but may not reflect consumers' actual program choices.

[7]The Standard Cotton programs are distinct from the regular Cotton programs, and both are available on the machine's control panel.

efficiency scale, ranging from A^{+++} to D, with a new scale ranging from A to G. Appendix A.1 provides the detailed formulas and classification thresholds under both regimes.

Two distinctions matter. First, neither the differences across wash programs nor the change in the Eco program across regimes imply fundamentally different engineering technologies. Wash programs generally share the same hardware but use different combinations of temperature, duration, water use, and mechanical action to serve different washing needs.[8] Temperature designations such as 40°C and 60°C are nominal program labels rather than mandated wash temperatures, and firms retain discretion over program inputs subject to regulatory constraints.[9] Second, energy classes are not directly comparable across regimes because both the test-load conditions and the EEI formula changed. A change in the displayed class therefore does not necessarily reflect a change in actual energy consumption. As the European Commission notes, a machine previously rated A^{+++} could receive a C rating under the new label without any change to the machine (European Commission, 2021).

Attribute Restrictions Beyond energy labeling, EU regulations impose attribute restrictions on the Eco program that do not apply to other wash programs. These restrictions constrain how firms balance cleaning performance against energy consumption. According to Sinner's Circle (Sinner, 1960), wash performance depends on four interacting factors: temperature, mechanical action, duration, and detergent chemistry.[10] Higher temperatures, longer cycles, and stronger mechanical action can improve stain and odor removal but may also increase fabric wear. Water serves as an intermediate input: more water facilitates rinsing but dilutes detergent and reduces friction between garments, while less water increases mechanical action but may hinder stain removal. Program inputs also affect electricity consumption, although their effects differ. Water heating accounts for more than 70 percent of electricity consumption in a typical wash cycle, making temperature reduction a primary means of reducing energy consumption. Lower temperatures, however, may compromise cleaning performance unless offset by adjustments in other inputs. Regulatory restrictions on program attributes can alter this tradeoff.

The main change in attribute restrictions under the 2021 reform was the introduction of duration limits on the Eco program. Under the earlier regime, manufacturers could improve measured energy efficiency by lowering wash temperatures and compensating with longer cycles while maintaining wash performance (Alborzi et al., 2017). However, excessively long cycles may discourage consumers

[8] For example, Cotton 60°C is intended for more intensive cleaning, Easy Care prioritizes gentler fabric treatment, and Quick Wash is designed for time convenience.

[9] In practice, Eco programs typically operate at wash temperatures between 30°C and 45°C, below their nominal temperature designations.

[10] Because detergent chemistry is standardized in regulatory tests, I treat it as exogenous to manufacturers' program-design decisions.

from using the Eco program, limiting energy savings in actual use. To improve time convenience and encourage Eco use, the 2021 reform imposed capacity-dependent duration caps, with separate limits at full, half, and quarter loads. The effect on consumer demand, however, is ambiguous. Consumers who value time convenience may prefer shorter Eco cycles, while those who associate longer cycles with better cleaning performance may find them less attractive and switch to other programs. The resulting effect on actual energy consumption depends on both firms' program-design responses and consumers' program choices.

The EU also imposes water-consumption and minimum wash-performance requirements under both regimes. The water constraint primarily aims to conserve water. The 2021 reform revised its measurement and introduced a separate rinsing-effectiveness requirement to limit residual detergent in clothing, which may cause skin irritation and allergic reactions (European Commission, 2019b). Water use therefore reflects the need to satisfy rinsing requirements as well as considerations of cleaning performance and energy efficiency. Both regimes also require a minimum level of stain removal, measured by a washing-efficiency index using standardized stain strips relative to a reference machine. The minimum threshold remained unchanged across regimes, although testing conditions changed. Advances in washing technology have made it possible to meet this requirement even at low temperatures (Josephy et al., 2015). However, the index captures only one dimension of cleaning performance: removal of standardized stains. It does not directly measure odor removal or hygienic performance, such as the reduction of microbial contamination, which households may also value (Alborzi et al., 2017). Meeting the regulatory threshold therefore does not imply that programs deliver the same cleaning performance from consumers' perspective. The index is reported in the product information sheet but is not displayed on the energy label, and reported values cluster near the regulatory threshold. In the model, I use this index to represent a pass-or-fail certification constraint, while allowing households' expected cleaning performance to depend on program inputs such as temperature, duration, and water use. Appendix A.2 reports the detailed regulatory constraints.

2.2 Data

I combine three main data sources: a product-level sales panel from GfK, program-level consumption information collected from product manuals, and a household laundry-habits survey conducted in Germany by Norstat on my behalf.

Product-Level Sales and Characteristics My primary data are a monthly panel of washing-machine sales in Germany from 2018 to 2023, provided by GfK. The data cover more than 90 percent of sales through electrical retailers across six geographic regions. For each region-month-model obser-

vation, the data report unit sales, revenues, and product characteristics, including energy-efficiency class, capacity, maximum spin speed, spin-drying efficiency, and noise level. To construct monthly market shares, I divide model-level unit sales by the estimated number of potential purchasers in each region-month. I define market size as the number of households in each region multiplied by the national washing-machine ownership rate and a monthly replacement probability implied by an average product lifespan of 12 years. To account for income heterogeneity in product demand, I supplement the sales data with the Panel on Household Finances (PHF), conducted by the Deutsche Bundesbank. I use the PHF to construct the household income distribution for demand estimation.

I focus on seven firms operating eight brands: BSH (Bosch and Siemens), Electrolux (AEG), Samsung, Miele, LG, Whirlpool (Bauknecht), and Arçelik (Beko). These firms account for approximately 93 percent of sales. I restrict the sample to washing machines with commonly purchased capacities of 7, 8, 9, and 10 kilograms. I define a product line as a combination of firm, load capacity, and maximum spin speed (e.g., Miele, 8 kg, 1,400 rpm). I define a product as a combination of firm, load capacity, maximum spin speed, energy-efficiency class, and quarter of initial market appearance (e.g., Miele, 8 kg, 1,400 rpm, class A, 2022Q1). The average transaction price is calculated as total revenue divided by total unit sales, aggregated across individual models within each product and region-month. The final sample contains 42 product lines and 349 products.

The composition of product introductions changes substantially over the sample period. Figure B10 in the appendix shows that early introductions are concentrated in lower energy-efficiency classes under the new label, while class A becomes increasingly prevalent over time. Exits are relatively infrequent, suggesting that the shift toward higher-rated products occurs primarily through entry rather than the withdrawal of lower-rated products.[11]

Program-Level Attribute Data My second data source consists of model-program-level information collected from 2,087 washing-machine user manuals. These manuals report energy consumption, load size, water consumption, duration, and temperature for individual wash programs. Unlike the energy label, which bases efficiency ratings on Eco performance, the manuals also report attributes of non-Eco programs, allowing me to compare regulated and unregulated programs within the same machine. The completeness of reporting differs across regulatory regimes. Under the 2021 regime, manufacturers are required to report Eco attributes at full, half, and quarter loads, including energy consumption, water use, duration, and temperature. Manuals also report commonly used non-Eco programs, including Cotton 20°C, Cotton 60°C, and Easy Care, while Cotton 40°C is reported by some brands. Under the earlier regime, reporting was less standardized, with temperature and other attributes frequently missing.

[11] Exit dates are not explicitly recorded. I identify an exit when a product disappears from the dataset or records zero sales for at least one quarter. Exit measurement is therefore right-censored at the end of the sample.

To improve coverage, I supplement the manuals with the European Product Registry for Energy Labelling (EPREL), using overlapping observations to validate reported attributes and fill missing values.[12] I impute the remaining missing attributes using the procedure described in Appendix B, obtaining complete program-level attributes for all models in the demand sample. I then aggregate these attributes to the product level by taking the unweighted average of each program's attributes across models belonging to the same product, as defined above.

Household Laundry-Habits Survey My third data source is a household laundry-habits survey that I commissioned in Germany in December 2025 and fielded through Norstat. The survey contains 508 completed responses from households that purchased washing machines under either the old or new labeling regime. I collect household demographics, appliance purchase history, and the model identifier and purchase date of each respondent's washing machine. To measure consumer information, I ask whether respondents considered the energy-efficiency class when purchasing their machine and whether they interpret the rating as reflecting the efficiency of specific wash programs or the entire machine. I also collect self-reported washing behavior, including the frequency of use of each available wash program and the frequency of washing lightly and heavily soiled laundry.[13]

Using the reported model identifiers, I match each respondent to the corresponding product manual and recover the attributes of the wash programs available on their machine. The matched data combine household characteristics, beliefs about energy labels, program-use frequencies, and program-specific energy consumption, temperature, water use, and duration.

Table 1 reports summary statistics for product characteristics, demand, and wash-program attributes. Panel A summarizes product characteristics and sales by energy-efficiency class. Higher-rated products tend to be more expensive and have larger capacities. Under the new label, average prices increase from €409 for class D to €601 for class A, while average capacity rises from 7.73 to 8.51 kg. Panel B reports program attributes and self-reported usage. Eco consumes less energy than Cotton 40°C and Cotton 60°C, averaging 0.87 kWh per cycle compared with 1.04 and 1.49 kWh, respectively. Eco also operates at lower temperatures and uses less water but runs substantially longer, averaging 225 minutes compared with approximately 184 minutes for the two Cotton programs.

[12] EPREL is the EU product database containing technical information submitted by manufacturers for energy-labelled appliances. Under the 2021 regime, its technical documentation includes Eco energy consumption, water use, duration, and temperature at full, half, and quarter loads.

[13] The English version of the questionnaire is provided in Appendix C.

Table 1: Summary Statistics

Panel A. Products					
Energy class	A^{+++}	D	C	B	A
Average price (€)	616.41	408.94	512.27	522.63	600.95
Average monthly unit sales	839.84	562.75	673.09	640.14	809.89
Average capacity (kg)	7.99	7.73	8.36	8.05	8.51
Number of products	139	44	40	41	85

Panel B. Program attributes and usage						
Program	Energy (kWh)	Water (L)	Duration (min)	Load size (kg)	Temperature (°C)	Reported cycles
Eco	0.87	56.38	224.53	8.12	37.98	7,304
Cotton 20°C	0.34	81.47	171.31	8.12	21.62	3,247
Cotton 40°C	1.04	83.00	184.30	8.12	41.26	17,317
Cotton 60°C	1.49	81.94	183.70	8.12	56.12	10,580
Easy Care	0.66	57.82	132.57	3.77	39.67	14,732

Notes: Panel A reports unweighted product-month means and distinct product counts for Germany, January 2018–December 2023. Prices are CPI-deflated euros. A^{+++} denotes the old label; D–A denote the new label. Panel B reports unweighted product means pooled across energy classes, including imputed attributes and excluding missing values. Eco combines old-label Standard 60 and new-label Eco at full load. Reported cycles pool both survey waves and cover all reported loads and products.

3 Regulation-Induced Energy Shifting

This section develops a stylized model, in the spirit of Holmstrom and Milgrom (1991) and Baker (1992), to illustrate how regulation targeting Eco can shift energy consumption toward unregulated programs, potentially increase average electricity consumption per cycle, and generate socially inefficient choices.

Consider a firm offering a washing machine with two programs, Eco and Cotton, with durations d_E and d_C. Holding temperature and other program inputs fixed, their energy consumption is given by $e_E(d_E)$ and $e_C(d_C)$, respectively. Both functions are increasing in duration, while Eco consumes less energy at a given duration. Consumers observe Eco energy consumption at purchase but mistakenly perceive it as representative of both programs. Because perceived energy costs are identical across programs, they enter product-purchase demand but cancel from program-choice comparisons. Program utility is therefore $v_r = \beta d_r + \varepsilon_r$, where β captures the average preference for duration and ε_r follows a standard type-I extreme-value distribution. Expected program utility

is

$$\mathcal{I}(d_E, d_C) = \log\left(e^{\beta d_E} + e^{\beta d_C}\right).$$

Taking product price P as given and normalizing the firm's markup to one, the firm chooses program durations to maximize demand net of separable design costs:

$$\max_{d_E, d_C} \quad Q\left(P + \kappa e_E(d_E) - \theta\mathcal{I}(d_E, d_C)\right) - C_E(d_E) - C_C(d_C) \qquad \text{s.t.} \qquad d_E \le \bar{d}_E,$$

where $Q' < 0$, $\kappa > 0$ captures consumers' valuation of perceived energy consumption, $\theta > 0$ measures the contribution of expected program utility to product demand, and C_E and C_C are separable design costs.[14]

Program designs are substitutes in expected utility, since $\frac{\partial^2 \mathcal{I}}{\partial d_E \partial d_C} = -\beta^2 s_E s_C \le 0$, where s_r denotes the program-choice probability. Shortening one program increases the marginal utility of extending the other. For $\beta > 0$, this means that extending Cotton becomes more attractive; for $\beta < 0$, it becomes less detrimental to utility. The cleaning-performance mechanism emphasized in this paper corresponds to $\beta > 0$. When this substitution carries over to product demand, tightening the Eco-duration cap increases the marginal return to extending Cotton. This requires that demand curvature not overturn the substitution in expected utility, as formalized by the cross-partial condition below.

Proposition 1 (Regulation-Induced Energy Shifting)**.** *Suppose the Eco-duration cap binds and Cotton duration has an interior local optimum satisfying the strict second-order condition. If* $\frac{\partial^2 Q}{\partial d_C \partial d_E} < 0$*, tightening the Eco-duration cap increases optimal Cotton duration and energy consumption locally:*

$$\frac{\partial d_C^*}{\partial \bar{d}_E} < 0, \qquad \frac{\partial e_C(d_C^*)}{\partial \bar{d}_E} < 0.$$

Proof. See Appendix D.1. □

The proposition establishes that restricting Eco duration can induce firms to increase energy consumption in the unregulated Cotton program. Whether this shifting raises average electricity consumption also depends on consumer program choices. Conditional on machine purchase, average electricity consumption per cycle is $\mathcal{E} = s_E e_E(d_E^*) + s_C e_C(d_C^*)$, where $s_E + s_C = 1$. With a binding Eco-duration cap, differentiating with respect to the cap yields

$$\frac{\partial \mathcal{E}}{\partial \bar{d}_E} = \underbrace{s_E \frac{\partial e_E}{\partial d_E}}_{\text{Direct Eco effect}} + \underbrace{s_C \frac{\partial e_C}{\partial d_C}\frac{\partial d_C^*}{\partial \bar{d}_E}}_{\text{Firm design response}} + \underbrace{\left[e_C(d_C^*) - e_E(d_E^*)\right]\frac{\partial s_C}{\partial \bar{d}_E}}_{\text{Consumer substitution}}.$$

[14] The reduced-form demand specification follows from logit product choice when perceived energy consumption and expected program utility enter purchase utility additively, with coefficients expressed in price-equivalent units.

Tightening the cap reduces Eco energy consumption but increases Cotton energy consumption. If $\beta > 0$, tightening the cap also shifts program use toward Cotton. When Cotton consumes more energy than Eco at the chosen durations, this substitution reinforces the increase in energy consumption. Average electricity consumption per cycle can therefore increase if the Cotton energy and program-substitution effects outweigh the Eco savings. The effect on total electricity use additionally depends on purchase quantities and washing frequency.

The interaction between energy-efficiency and duration constraints can further affect firms' design responses. When firms can adjust other Eco inputs, such as temperature, a tighter duration cap creates room under a binding Eco energy target to increase these inputs, potentially offsetting the shift toward Cotton. A stricter energy target limits this adjustment and can strengthen incentives to improve Cotton instead. Appendix D.2 formalizes this interaction.

These design responses also have welfare implications when consumers have incomplete energy information. Consumers perceive Eco energy consumption as representative of both programs, overlooking differences in their actual electricity costs. A social planner instead accounts for program-specific energy consumption when choosing program designs and allocating washing cycles, alongside the benefits of cleaning performance and the costs of program design. This information gap can lead firms to choose excessively long and energy-intensive Cotton programs, although the comparison with the social optimum depends on these benefits and costs. Appendix D.3 provides the conditions for this result.

4 Descriptive Evidence

This section presents three empirical patterns that motivate the structural model in Section 5 and relate to the mechanisms developed in Section 3. First, following the reform, Eco energy consumption per cycle, while Cotton energy consumption increases and remains elevated, documenting divergent responses across programs. Second, products more exposed to the Eco-duration cap exhibit longer and more energy-intensive Cotton cycles relative to old-label benchmarks, consistent with the proposed mechanism of regulation-induced energy shifting. Third, consumers respond to energy ratings at purchase but often interpret Eco-based ratings as measuring whole-machine efficiency, even though non-Eco programs account for most reported washing cycles.

4.1 Changes in Program-Level Energy Consumption

Because program characteristics are fixed over a product's lifetime, I examine changes in mean energy consumption across products offered within each product line, weighting products equally. These changes reflect product turnover rather than shifts in sales shares. I restrict the sample

to product lines observed in every quarter from 2018Q1 to 2023Q4. For comparability across labeling regimes, I measure energy consumption per kilogram and apply common load weights to Eco programs.[15] I estimate

$$\log \bar{e}_{\ell rt} = \sum_{r' \in \mathcal{R}} \sum_{k \neq 2019Q3} \beta_{kr'} \mathbf{1}\{t = k\}\mathbf{1}\{r = r'\} + \mu_{\ell r} + \mu_t + \varepsilon_{\ell rt}, \tag{1}$$

where $\bar{e}_{\ell rt}$ is mean energy consumption per kilogram in product line ℓ, program r, and quarter t. The specification includes product-line-by-program fixed effects, $\mu_{\ell r}$, and quarter fixed effects, μ_t. The set $\mathcal{R}$ contains Eco, Cotton 40°C, Cotton 60°C, and Easy Care. Wool serves as the reference program because its specialized requirements for delicate fabrics make its design less likely to respond to the reform. The coefficients measure changes relative to Wool, with 2019Q3 as the reference quarter. Standard errors are clustered by product line.

Figure 1 presents the results for Eco, Cotton 60°C, Cotton 40°C, and Easy Care. Eco energy consumption initially rises before declining after 2021, whereas consumption in the three unregulated programs rises and remains elevated. By 2023Q4, the estimated changes relative to Wool and the reference quarter are −8.6 percent for Eco, 14.9 percent for Cotton 60°C, 8.4 percent for Cotton 40°C, and 7.9 percent for Easy Care. These patterns indicate that improvements in regulated Eco efficiency do not extend to other commonly used programs.

Figure E2 in the appendix reports two robustness checks using alternative measures of Eco energy consumption. The alternative old-program definition yields a similar trend when common load weights are applied across regimes.[16] By contrast, label-reported Eco consumption shows a much larger decline under the respective pre- and post-reform testing protocols, indicating that part of the apparent improvement reflects changes in measurement.

4.2 Accounting for Changes in Energy Consumption

To understand the inputs underlying changes in energy consumption, I separate changes in program attributes from persistent differences in conditional energy efficiency across products. Following the distinction between attribute choices and technological progress in Knittel (2011), I estimate

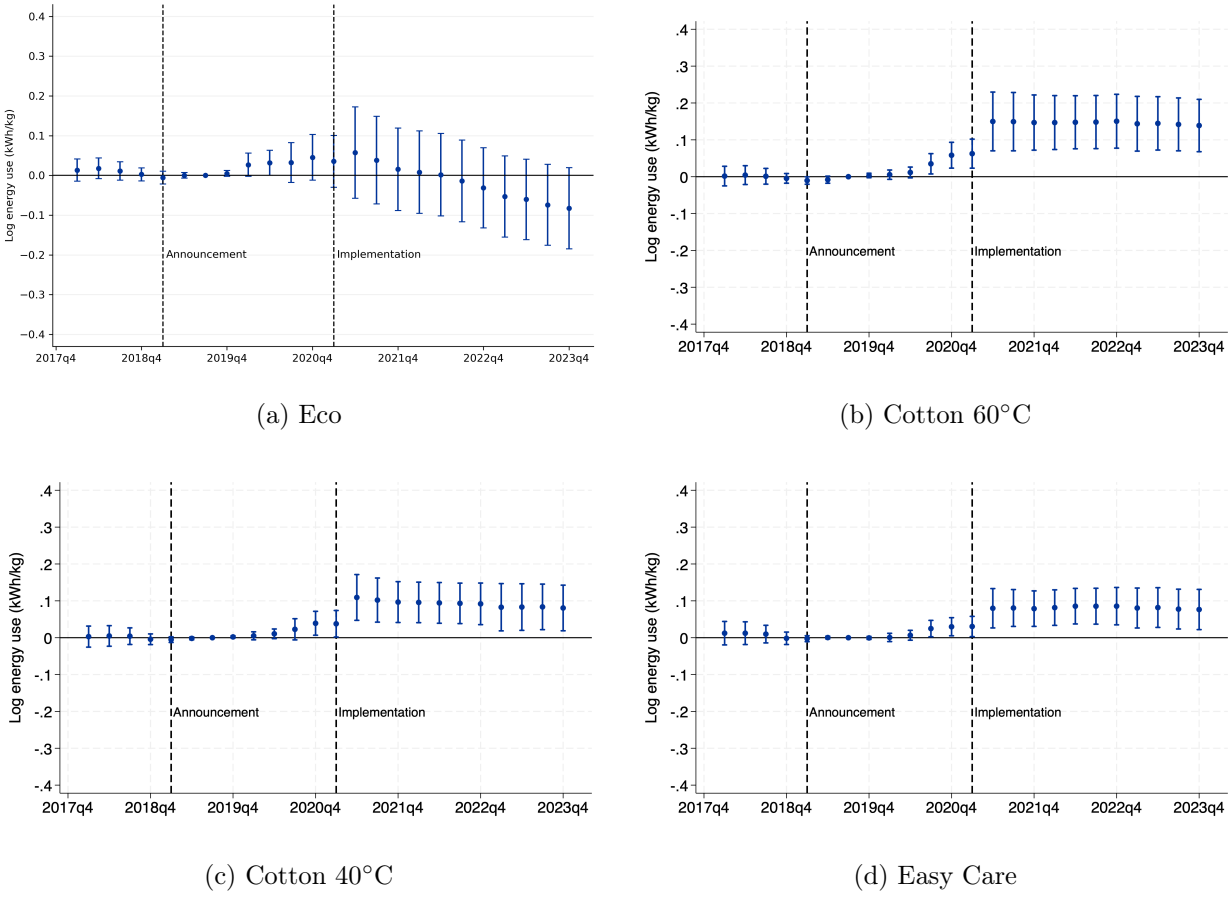
$$\log e_{jro} = z_j + \delta_{g(r)} + \theta_T \log T_{jro} + \theta_L \log L_{jro} + \theta_d \log d_{jro} + \theta_w \log w_{jro} + \varepsilon_{jro}, \tag{2}$$

where e_{jro} is electricity consumption per cycle for product j, program r, and load condition o. The inputs are temperature T, laundry load L, duration d, and water use w. Product fixed effects z_j

[15] I apply the new-label weights, normalized across full and half load, to Eco 40–60 under the new regime and to Standard Cotton 60°C at full load and Standard Cotton 60°C at half load under the old regime.

[16] The alternative measure combines Standard Cotton 60°C and Standard Cotton 40°C under the old regime.

Figure 1: Changes in Program-Level Energy Consumption



(a) Eco

(b) Cotton 60°C

(c) Cotton 40°C

(d) Easy Care

Notes: Panels (a)–(d) report changes in log energy consumption per kilogram for Eco, Cotton 60°C, Cotton 40°C, and Easy Care, respectively, relative to Wool. Dots denote the estimated coefficients $\beta_{kr'}$ from equation (1), normalized to zero in 2019Q3. The sample contains 31 product lines (firm × capacity × spin speed), observed in every quarter from 2018Q1 to 2023Q4, with products weighted equally within each line. Bars denote 95 percent confidence intervals based on standard errors clustered by product line. Dashed vertical lines mark the reform's announcement in March 2019 and implementation in March 2021.

capture persistent differences in energy consumption conditional on these inputs, while program-group effects $\delta_{g(r)}$ distinguish Cotton and Easy Care from Eco. The input coefficients are estimated using variation across programs and load conditions within products.[17]

The fixed effects may absorb hardware efficiency, unobserved operating characteristics such as mechanical action, and systematic measurement differences. I therefore interpret them as conditional energy efficiency rather than as identified technology investments. The structural counterfactuals hold the product-specific component fixed while allowing firms to adjust the explicitly modeled

[17]Unlike the year effects used to capture changes in the technological frontier in Knittel (2011), product fixed effects exploit the availability of multiple program and load observations for each machine.

program inputs.

Appendix Table A1 reports the estimates and a robustness specification restricted to old and new Eco observations. All four input coefficients are positive, with temperature having the largest estimated elasticity. The main specification estimates elasticities of 1.580 for temperature, 0.401 for duration, and 0.321 for water use. The Eco-only specification preserves these signs, although the duration coefficient is smaller. Sample construction, estimation details, and the contribution calculations are described in the appendix.

Table A2 relates changes in energy consumption per kilogram to changes in program inputs. Products receive equal weights, so the comparisons reflect turnover in the offered portfolio rather than changes in sales shares. A central pattern is the divergence in cycle duration. By 2023Q4, relative to the reference quarter specified in the table, Eco duration falls by 8.0 percent, while Cotton 60°C, Cotton 40°C, and Easy Care durations rise by 18.2, 17.9, and 15.6 percent, respectively. Given the positive estimated duration coefficient, these changes contribute to lower Eco electricity consumption and higher consumption in unregulated programs, holding other inputs fixed.

Duration is also quantitatively important in the contribution calculations. Longer cycles are the largest positive input contribution for both Cotton programs, accounting for 29.2 percent of absolute fitted contributions for Cotton 60°C and 31.0 percent for Cotton 40°C. Greater water use further raises Cotton consumption, while higher temperature is the largest positive contribution for Easy Care. For Eco, shorter cycles and lower temperatures contribute to energy savings, partly offset by increased water use. These shares describe the relative magnitudes of fitted contributions, including the unreported conditional-efficiency component, rather than fractions of the observed energy change.

These patterns highlight duration as an important margin underlying the divergence between regulated and unregulated energy consumption. They motivate modeling firms' duration choices jointly with other program inputs to examine how a cap on Eco duration can reshape the design and energy consumption of unregulated alternatives.

4.3 Eco Duration-Cap Exposure and Non-Eco Program Design

Proposition 1 in Section 3 predicts that a binding Eco duration cap shifts performance-related inputs toward unregulated Cotton programs. To investigate this mechanism, I examine whether products more exposed to the cap have longer and more energy-intensive Cotton programs relative to pre-policy benchmarks.

The sample consists of dual-label products with observed old- and new-label energy classes. Because their non-Eco outcomes are not observed under both regimes, I compare them with matched

Table 2: Duration-Cap Exposure and Cotton and Easy Care Program Design

	Cotton 60°C			Cotton 40°C			Easy Care		
	Energy	Duration	Water	Energy	Duration	Water	Energy	Duration	Water
Exposure	0.570***	0.211***	-0.061	0.212***	0.096*	-0.051	0.136*	-0.072	-0.153***
	(0.055)	(0.058)	(0.045)	(0.061)	(0.050)	(0.055)	(0.081)	(0.058)	(0.053)
Capacity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Brand FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	223	223	223	223	223	223	223	223	223
R-squared	0.554	0.307	0.460	0.373	0.487	0.408	0.111	0.574	0.419

Notes: The dependent variables are log differences between each dual-label product's energy, duration, or water consumption and the corresponding mean among old-label-only donors. Donors match the same brand and exact Standard 60 full-load duration and loading. Current completed inputs, including imputed values, are used. Exposure is max(0, Standard full-load duration minus the Eco full-load cap), divided by that cap. All regressions include capacity and brand fixed effects. * p<0.10, ** p<0.05, *** p<0.01.

old-label-only A^{+++} products. Appendix E.2 describes the matching procedure. For each program r and outcome $q \in \{e, d, w\}$, denoting energy consumption, duration, and water use, respectively, I estimate

$$\log q_{jr}^{\text{new}} - \log \widehat{q}_{jr}^{\text{old}} = \beta_r^q \, \text{Exposure}_j + \mu_{k(j)}^{q,r} + \eta_{b(j)}^{q,r} + \varepsilon_{jr}^q, \tag{3}$$

where $\widehat{q}_{jr}^{\text{old}}$ is the mean outcome among matched old-label products, and $\mu_{k(j)}^{q,r}$ and $\eta_{b(j)}^{q,r}$ denote capacity and brand fixed effects. Exposure measures how much pre-policy regulated-program duration exceeds the new cap, capturing potential pressure to shorten Eco rather than whether the constraint binds for the redesigned program:

$$\text{Exposure}_j = \max\left\{0, \frac{\overline{d}_{f(j)k(j),\text{reg}}^{\text{pre}} - ub_{k(j)}}{ub_{k(j)}}\right\}, \tag{4}$$

where $\overline{d}_{f(j)k(j),\text{reg}}^{\text{pre}}$ is the mean pre-policy duration within a firm–capacity group and $ub_{k(j)}$ is the new duration cap. Identification comes from variation in exposure across firm–capacity groups, conditional on brand and capacity fixed effects.

Table 2 reports the estimates. A 10-percentage-point increase in exposure is associated with 5.9 percent higher energy consumption and 2.1 percent longer duration in Cotton 60°C relative to old-label benchmarks. Cotton 40°C energy consumption increases by 2.1 percent and duration by 1.0 percent, although the latter is significant only at the 10 percent level. Easy-Care energy consumption increases by 1.4 percent, also significant at the 10 percent level, while its duration response is not statistically significant. Water use does not change significantly in either Cotton program but declines by 1.5 percent in Easy-Care. These findings are consistent with Proposition 1,

suggesting that greater pressure to shorten Eco is associated with longer and more energy-intensive Cotton programs, particularly Cotton programs.

4.4 Consumer Program Use, Label Beliefs, and Purchase Responses

The household survey reveals a substantial gap between how consumers use washing machines and what they believe energy ratings measure. Figure E3 shows that Eco accounts for only 9% of reported washing cycles among new-label households, compared with 25% for Cotton 40°C, 17% for Cotton 60°C, and 22% for Easy Care. Eco usage is also limited among old-label households, accounting for 13% of cycles. Despite this limited use, Figure E4 shows that 71% of new-label households mistakenly believe the energy class reflects energy consumption across all programs, even though it is determined by Eco performance alone. Moreover, 67% rely on the energy class rather than reported kWh consumption to assess energy costs. These findings suggest that consumers may value energy ratings as indicators of whole-machine efficiency, even though the ratings do not reflect energy consumption in their most frequently used programs.

To investigate purchase responses, I exploit the 2021 label rescaling from A^{+++}–D to A–G. Before the reform, most products were rated A^{+++}, providing little differentiation across models. The rescaling redistributed these products across the new A–D classes, creating variation in the information conveyed by energy ratings. I examine whether products receiving more favorable ratings experienced larger sales gains when the new labels became visible at retail.

Importantly, the transition in product design preceded the retail label switch. Beginning in 2020, manufacturers started updating control panels and Eco programs to accommodate the new regulatory requirements, following permission to display the Eco designation from late December 2019.[18] The new energy labels, however, became visible at retail only on March 1, 2021. Thus, for products whose designs remained unchanged around the rollout, the label switch altered the information presented to consumers without changing their underlying energy consumption. This timing allows me to examine the demand response to the newly displayed energy ratings separately from the earlier adjustments in program design.

Using a pre-implementation snapshot, I predict each product's energy class under the new system and define $E_j = 1$ for products predicted to receive A, B, or C, and $E_j = 0$ for those predicted to receive D. I estimate an event study centered on the March 2021 retail rollout, using February 2021 as the reference month:

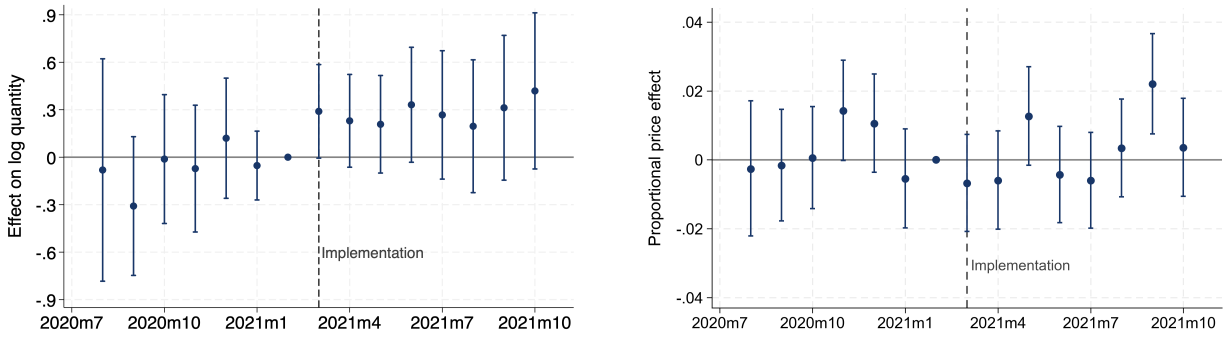
$$\log(q_{jmt}) = \sum_{k \neq -1} \beta_k \mathbf{1}\{\tau_t = k\} E_j + \beta_p \log(p_{jmt}) + \mu_{jm} + \mu_t + \epsilon_{jmt}, \tag{5}$$

[18] See Commission Regulation (EU) 2019/2023, Article 11, and Czyzewski (2022).

where q_{jmt} and p_{jmt} denote product sales and prices, τ_t is the number of months relative to March 2021, μ_{jm} are product–market fixed effects, and μ_t are calendar-month fixed effects. I censor each product's panel at its first observed post-rollout rating change to limit contamination from subsequent redesign. The coefficients β_k capture differential sales changes between products predicted to receive A–C and those predicted to receive D, holding product identity fixed.

Figure 2 reports the estimates. Pre-reform coefficients are close to zero, indicating little evidence of differential pre-trends. Following the retail rollout, products predicted to receive A–C experience an immediate and persistent increase in sales relative to D products. The average post-reform coefficient of 0.44 implies approximately 55% higher relative sales. Because product-design adjustments began before the retail rollout, the sharp increase in relative sales around March 2021 is consistent with a demand response to the newly salient energy ratings rather than contemporaneous changes in energy performance. Together with the survey evidence, these results suggest that consumers place substantial value on energy ratings, which they often mistakenly interpret as measuring whole-machine efficiency. The corresponding price estimates show little evidence of differential price changes following the rollout.

Figure 2: Demand and Price Effects



(a) Demand effect

(b) Price effect

Notes: Event-study estimates around the March 2021 energy-label rescaling. Panel (a) plots the coefficients $\widehat{\beta}_k$ from equation (5) for log product sales, controlling for log price. Panel (b) reports price effects from $\log p_{jmt} = \sum_{k \neq -1} \beta_k^p \mathbf{1}\{\tau_t = k\} E_j + \mu_{jm} + \mu_t + \varepsilon_{jmt}^p$. Both specifications include product–market and calendar-month fixed effects and omit February 2021. Points are estimates and bars are 95 percent confidence intervals. Both regressions cluster standard errors by product. Dashed lines mark March 2021, when the new energy labels became visible to consumers.

5 Model

I develop a static equilibrium model of washing-machine demand and supply to examine how energy labels and ecodesign requirements affect product design, consumer choices, and energy consumption.

Within each period, the model follows four stages:

S1. Product Entry. Private portfolio-level entry shocks are realized. Firms simultaneously choose which product configurations to introduce, taking the distribution of rivals' entry decisions as given.[19]

S2. Program Design. Market entry outcomes are realized and observed by all firms. Firms choose program inputs when introducing an energy class not previously offered on a product line, relative to predetermined reference designs and subject to energy-efficiency and duration constraints. Repeated introductions of a previously offered class inherit its existing program design. Program inputs remain fixed over the product's lifetime.

S3. Price Setting and Product Choice. Product-level demand and marginal-cost shocks are realized. Firms set prices in Bertrand–Nash equilibrium, and households choose among the available washing machines and the outside option. Product demand internalizes the expected utility from subsequent program use.

S4. Program Use. Program-use shocks are realized before each wash occasion. Households then choose among the available wash programs of the purchased washing machine.

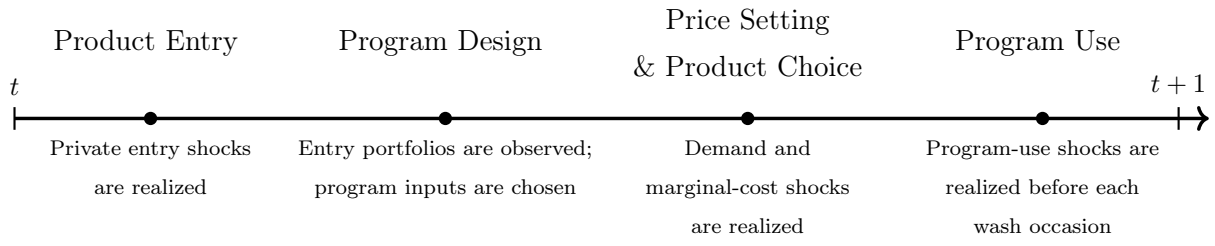


Figure 3: Timing within period t.

5.1 Demand

A market is indexed by (m, t), where m denotes region and t quarter. I model household demand as a two-stage decision problem, combining a differentiated-products demand system in the spirit of Berry et al. (1995) with a subsequent usage decision, as in the appliance ownership and utilization framework of Dubin and McFadden (1984).

There is a continuum of households indexed by i. Household heterogeneity consists of income y_i, the household's share of high-stain laundry τ_i, and persistent preference draws for program

[19] As shown in Figure B10, exit among new-label products is rare and concentrated near the end of the sample. I abstract from endogenous exit decisions and also treat the exit of legacy products as exogenous.

temperature and duration. Income shifts price sensitivity, while the high-stain share and preference draws shift how households value wash-program inputs. I begin with the program-use decision because it determines the continuation value associated with owning each product.

Program Use Let $r \in \mathcal{R}$ index a common set of inside wash programs: Cotton 20°C, Cotton 40°C, Cotton 60°C, Easy Care, and Eco. Let $r = 0$ denote an outside option, interpreted as the use of any other program, and let s index wash occasions. Conditional on owning product j, household i's utility from program $r \in \mathcal{R}$ on occasion s is

$$v_{ijrs} = \delta_{c(r)} + \rho_i^T T_{jr} + \rho_i^d d_{jr} + \rho_w w_{jr} + \rho_L L_{jr} + \varepsilon_{ijrs}^v, \tag{6}$$

where T_{jr}, d_{jr}, w_{jr}, and L_{jr} denote temperature, duration, water use, and load size.[20] The shock ε_{ijrs}^v is i.i.d. type-I extreme value and is redrawn across wash occasions; the outside option is normalized to $v_{ij0s} = \varepsilon_{ij0s}^v$. Preferences for temperature and duration vary with the household's share of high-stain laundry according to $\rho_i^q = \rho_0^q + \rho_1^q \tau_i + \sigma_q \nu_i^q$, for $q \in \{T, d\}$ and $\nu_i^q \sim N(0,1)$. These preference draws remain fixed across wash occasions. The program-category fixed effect $\delta_{c(r)}$ captures baseline preferences and washing habits associated with category $c(r)$.[21]

Let $\bar{v}_{ijr} \equiv v_{ijrs} - \varepsilon_{ijrs}^v$ denote the deterministic component of utility from program r. Under the type-I extreme value assumption, the conditional probability that household i chooses program $r \in \mathcal{R}$ on product j is

$$P_{ijr} = \frac{\exp\left(\bar{v}_{ijr}\right)}{1 + \sum_{r' \in \mathcal{R}} \exp\left(\bar{v}_{ijr'}\right)}. \tag{7}$$

The expected utility from product j's program bundle is

$$\mathcal{I}_{ij} \equiv \mathbb{E}_{\varepsilon^v}\left[\max_{r \in \{0\} \cup \mathcal{R}} v_{ijrs}\right] = \log\left(1 + \sum_{r \in \mathcal{R}} \exp\left(\bar{v}_{ijr}\right)\right), \tag{8}$$

up to the Euler–Mascheroni constant, which is common across products and normalized away.

[20] Households may not observe individual program inputs directly, but they perceive the resulting washing quality. I represent perceived quality as a function of the underlying inputs without requiring households to evaluate each input separately. Persistent product-level characteristics may also affect perceived washing quality. I assume that their common contribution across all programs, including the outside program option, enters product-choice utility.

[21] Program categories are Cotton, Easy Care, and Eco. In particular, Cotton 20°C, Cotton 40°C, and Cotton 60°C share the same category fixed effect.

Product Choice In each market (m,t), household i chooses among the set of available washing machines $\mathcal{J}_{mt}$ and an outside option. For product $j \in \mathcal{J}_{mt}$, indirect utility is

$$\begin{aligned} u_{ijmt} = & -\alpha_i p_{jmt} + \sum_{c \in \mathcal{C}} \beta_c \mathbf{1}\{\text{ec21}(j) = c\}\mathbf{1}_{\text{new},t} + \beta^I \mathcal{I}_{ij} + x_j'\beta^x \\ & + \delta_{f(j)} + \delta_{\text{cohort}(j)} + \delta_m + \delta_t + \xi_{jmt} + \epsilon_{ijmt}, \end{aligned} \tag{9}$$

where ϵ_{ijmt} is i.i.d. type-I extreme value and the outside-option mean utility is normalized to zero. Price sensitivity varies with household income according to $\alpha_i = \exp(\bar{\alpha} + \alpha^y \log y_i)$.

The indicator $\mathbf{1}_{\text{new},t}$ equals one under the new labeling regime, while $\mathbf{1}\{\text{ec21}(j) = c\}$ indicates whether product j belongs to energy class $c \in \mathcal{C} = \{A, B, C, D\}$, with class D normalized to zero. I model the displayed energy class as a coarse signal of energy efficiency that enters utility directly, without separately specifying consumers' perceived product-specific energy costs. This specification accommodates responses to energy certifications without requiring consumers to fully process the underlying energy-cost information (Houde, 2018). The coefficients β_c therefore capture consumers' valuation of each displayed energy class under the new label.[22] I omit the pre-reform energy-class effects because almost all products were rated A^{+++} before the reform.

The inclusive value $\mathcal{I}_{ij}$ links program use to product choice by summarizing household i's expected value from product j's program bundle. Although I do not explicitly model reputation formation, this provides a parsimonious way to capture a reputation-like channel: consumers may learn about downstream product performance through reviews and other pre-purchase information (Reimers and Waldfogel, 2021), giving firms an incentive to design programs that generate higher expected use value. The coefficient β^I determines how strongly this expected program-use value enters the product-purchase decision.

The vector x_j contains predetermined, time-invariant product characteristics, including capacity indicators, spin speed, and the product-specific conditional-efficiency component z_j from equation (2). As discussed in Section 4.2, this component may capture hardware efficiency, omitted operating characteristics, and systematic measurement differences; it is not an identified measure of technology investment. I hold it fixed in firms' design decisions and in the counterfactuals. The terms $\delta_{f(j)}$, $\delta_{\text{cohort}(j)}$, δ_m, and δ_t denote firm, entry-cohort, region, and quarter fixed effects, respectively. Entry cohorts are defined by the quarter of initial market appearance.[23] The term ξ_{jmt} captures product-market-time attributes observed by households but not by the econometrician.

[22] Eco-program energy consumption bunches near class thresholds, as shown in Figure A7, so energy-class and capacity fixed effects absorb much of the variation in energy consumption displayed on the label. This leaves limited variation to separately identify sensitivity to energy costs. I therefore use class premiums to capture demand for energy efficiency, including the energy-cost savings consumers associate with each class.

[23] BSH is the only firm represented by two brands in the sample, Bosch and Siemens. Given their common ownership and closely related hardware and program designs, I combine them into a single firm.

Let $\bar{u}_{ijmt} = u_{ijmt} - \epsilon_{ijmt}$ denote the deterministic component of product utility. Conditional on household characteristics, the probability that household i purchases product j is

$$s_{ijmt} = \frac{\exp\left(\bar{u}_{ijmt}\right)}{1 + \sum_{k \in \mathcal{J}_{mt}} \exp\left(\bar{u}_{ikmt}\right)}. \tag{10}$$

The market share of product j integrates this choice probability over the distribution of household heterogeneity:

$$s_{jmt} = \int s_{ijmt}\, dF_{mt}(y_i, \tau_i, \nu_i^T, \nu_i^d). \tag{11}$$

5.2 Supply

Actions. Let $\mathcal{L}_{ft}$ denote firm f's set of potential product lines in quarter t. Each line is defined by firm × capacity × spin speed. On each line $\ell \in \mathcal{L}_{ft}$, the firm chooses $a_{f\ell t} \in \mathcal{A} \equiv \{0, A, B, C, D\}$, where 0 denotes no entry and $c \in \{A, B, C, D\}$ denotes introduction of a product in energy class c. Each line therefore permits at most one introduction per quarter. The firm's portfolio action is $a_{ft} = (a_{f\ell t})_{\ell \in \mathcal{L}_{ft}} \in \mathcal{F}_{ft} \equiv \mathcal{A}^{|\mathcal{L}_{ft}|}$, and $a_t = (a_{ft}, a_{-f,t})$ denotes the joint entry action of all firms. For a product j introduced on line $\ell(j)$, its energy class is $c_j = a_{f\ell(j)t}$. Product lines also determine design-cost references. Let $\mathcal{J}_t(a_t)$ denote the resulting set of available products.

For each entering product j eligible for redesign, let q_{fjrt} denote the vector of chosen inputs for wash program r, consisting of temperature and duration where these are adjustable. Water use, load size, and product-specific conditional efficiency are predetermined. I collect the adjustable program inputs in q_{ft} and denote the industry-wide profile by $q_t = (q_{ft}, q_{-f,t})$. Program inputs of incumbent products and repeated introductions of previously offered classes remain fixed. Given the resulting product portfolio and program designs, firms then compete in prices across markets.

Price Setting. Given the realized product portfolio $\mathcal{J}_t(a_t)$ and program designs q_t, firms compete in prices across markets. Firm f chooses $\{p_{jmt}\}_{j \in \mathcal{J}_{ft}(a_t)}$ to maximize

$$\pi_{ft}\big(\mathcal{J}_t(a_t), q_t\big) = \sum_m \sum_{j \in \mathcal{J}_{ft}(a_t)} \big(p_{jmt} - mc_{jt}\big) s_{jmt}\big(p; \mathcal{J}_t(a_t), q_t\big) M_{mt}, \tag{12}$$

where s_{jmt} is product j's market share and M_{mt} is market size. Prices are determined in Bertrand–Nash equilibrium. Marginal cost is

$$\log mc_{jt} = x_j^{mc\prime}\gamma + \gamma_{f(j)} + \gamma_t + \omega_{jt}, \tag{13}$$

where x_j^{mc} contains capacity, spin speed, and product-specific conditional efficiency, $\gamma_{f(j)}$ and γ_t are firm and quarter fixed effects, and ω_{jt} is an idiosyncratic marginal-cost shock observed by firms.

The adjustable program inputs are control-algorithm parameters and therefore do not directly enter marginal production cost.

Program Design. After entry outcomes are realized and observed, firms simultaneously choose program inputs for products introducing an energy class not previously offered on their line. Let $\mathcal{J}_{ft}^{\text{red}}(a_{ft})$ denote this set of products. For product j entering on line $\ell(j)$, let $q_{f\ell(j)r}^{\text{ref}}$ denote its predetermined reference design. Duration references are anchored to the matched pre-reform A^{+++} benchmark, while temperature references depend on the line's prior design history; the detailed construction is described in Section 6. Repeated introductions of a previously offered class retain that class's first design. Define $\Delta q_{fjrt} = q_{fjrt} - q_{f\ell(j)r}^{\text{ref}}$. For $j \in \mathcal{J}_{ft}^{\text{red}}(a_{ft})$, the design cost is

$$DC_{fjt} = \sum_{r \in \mathcal{R}} \left[(\kappa_r^d)' \Delta q_{fjrt} + \Delta q_{fjrt}' C_r^d \Delta q_{fjrt} \right], \tag{14}$$

where κ_r^d and $C_r^d \succeq 0$ govern the linear and quadratic costs of departing from the reference design.

Conditional on the realized portfolio a_t, firms play a complete-information program-design game. Taking rivals' program inputs as given and holding inherited designs fixed, firm f solves

$$\begin{aligned} W_{ft}(a_t) = \max_{q_{ft}} \Bigg\{ & \mathbb{E}_{\xi,\omega} \left[\pi_{ft}\big(\mathcal{J}_t(a_t), q_t; \xi, \omega\big) \right] - \sum_{j \in \mathcal{J}_{ft}^{\text{red}}(a_{ft})} DC_{fjt} \Bigg\} \\ \text{s.t.} \quad & \text{EEI}_j(q_{fj,\text{Eco},t}) \in \mathcal{E}(c_j), \\ & d_{jo}(q_{fj,\text{Eco},t}) \le \overline{d}_o(k_j), \qquad o \in \{F, H, Q\}. \end{aligned} \tag{15}$$

Here k_j denotes product j's rated capacity, $\mathcal{E}(c_j)$ is the EEI range associated with the energy class chosen at entry, and $\overline{d}_o(\cdot)$ denotes the applicable load-specific duration limits. The constraints apply to products eligible for redesign, with predetermined inputs suppressed in the notation. The exact regulatory formulas are provided in the Appendix A. I hold water use fixed because it also reflects rinsing and other requirements, as discussed in Section 2.1. I assume that the minimum Eco washing-performance requirement is readily satisfied and does not bind over the designs considered. Non-Eco programs are unconstrained by the displayed requirements but affect program use, product demand, and operating profits. Evaluated at the program-design equilibrium, $W_{ft}(a_t)$ is the continuation payoff that firms anticipate at the entry stage.

Product Entry. Firms simultaneously choose whether to introduce a product on each potential line and, conditional on entry, its energy class, taking the distribution of rivals' entry decisions, $\mathcal{P}_{-f,t}$, as given. The vector $a_{ft} = (a_{f\ell t})_{\ell \in \mathcal{L}_{ft}}$ constitutes the firm's introduction portfolio, with feasible portfolios belonging to $\mathcal{F}_{ft}$.

Each firm privately observes portfolio-specific payoff shocks and solves

$$V_{ft} = \max_{a_{ft} \in \mathcal{F}_{ft}} \left\{ \beta_\pi \mathbb{E}_{\mathcal{P}_{-f,t}} \left[W_{ft}(a_{ft}, a_{-f,t}) \right] - C_{ft}^{\text{entry}}(a_{ft}) + \eta_{ft}(a_{ft}) \right\}, \tag{16}$$

where W_{ft} is the continuation payoff from the subsequent program-design equilibrium, comprising current-period operating profit net of design costs. The shocks $\eta_{ft}(a_{ft})$ are i.i.d. across firms and feasible portfolios with a type-I extreme-value distribution whose scale is normalized to one. The coefficient $\beta_\pi > 0$ converts monetary payoffs into normalized payoff units and governs firms' responsiveness to expected profitability.

I specify a common launch cost and a switching premium that varies by destination class and between the transition and post-reform periods:

$$C_{ft}^{\text{entry}}(a_{ft}) = \sum_{\substack{\ell \in \mathcal{L}_{ft} \\ a_{f\ell t} \neq 0}} \left[\kappa^e + \kappa^{e,\text{sw}}_{g(a_{f\ell t}),h(t)} \mathbf{1}\left\{ a_{f\ell t} \notin \mathcal{C}^{\text{last}}_{f\ell,t} \right\} \right]. \tag{17}$$

Here, $\mathcal{C}^{\text{last}}_{f\ell,t}$ denotes the set of energy classes introduced on line ℓ in its most recent introduction quarter before t. For a line with no previous introduction, I set $\mathcal{C}^{\text{last}}_{f\ell,t} = \varnothing$, so its first introduction incurs the corresponding switching premium. The index $h(t) \in \{\text{tran}, \text{post}\}$ distinguishes the transition period, 2020Q1–2020Q4, from the post-reform period, beginning in 2021Q1. The mapping $g(A) = A$, $g(B) = B$, and $g(C) = g(D) = CD$ pools switching premiums for classes C and D within each period. This pooling applies to the premium coefficients; switching between C and D still incurs the premium.

The common launch cost κ^e captures fixed expenses beyond the continuous design costs included in W_{ft} and is incurred for every introduction. Introducing a class represented in $\mathcal{C}^{\text{last}}_{f\ell,t}$ incurs only κ^e, while introducing a different class c additionally incurs $\kappa^{e,\text{sw}}_{g(c),h(t)}$. Returning to an older class also incurs this switching premium if that class was absent from the line's most recent introduction quarter, although it inherits its previously established program design.

5.3 Discussion

The model makes two main simplifying assumptions. First, I assume that firms' private payoff shocks are i.i.d. type-I extreme value across feasible portfolios. This assumption yields exact conditional-logit probabilities, allowing estimation and Gibbs sampling without enumerating the full portfolio choice set while preserving interactions among products in deterministic profits. Economically, portfolio-level shocks can represent private coordination costs, retailer agreements, or assessments of a product lineup. The restriction is their independence across overlapping portfolios, which rules out unobserved product-specific costs shared by portfolios containing the same product. With product-level shocks and interacting profits, conditioning on the firm's other chosen products generally

changes the distribution of the remaining product's shock. The conditional probabilities therefore need not retain the logit form required by the current estimation and sampling procedure. Even evaluating a single conditional probability can then be computationally demanding, because it requires integrating over shocks consistent with the firm's optimal choices across its portfolio.

Second, I treat consumer purchase and firm entry decisions as static. I assume that consumers replace washing machines only upon failure, rather than optimally timing replacement as in Goettler and Gordon (2011). Unlike microprocessors, where consumers may delay upgrades in anticipation of technological improvements and price declines, washing-machine replacement is more likely driven by necessity. Firms likewise maximize current-period profits, even though products remain on the market for approximately 2.5 years on average. Following Wollmann (2018), static entry decisions can approximate dynamic investment decisions when current profitability predicts future returns, with estimated entry costs reflecting an implicit hurdle rate. These assumptions abstract from dynamic replacement and entry incentives to focus on the paper's main mechanism: regulation-induced changes in program design and household substitution across wash programs.

6 Estimation

I first estimate program preferences, product demand, and the engineering relationship, then recover marginal costs from pricing conditions and estimate design and entry costs.

6.1 Program Use, Demand, and Marginal Cost

For survey household i, let n_{ir} count reported uses of program r. Simulated maximum likelihood integrates household-specific tastes jointly across its reported washes:

$$\max_{\rho} \sum_{i} \log \left[\frac{1}{R} \sum_{b=1}^{R} \prod_{r \in \{0\} \cup \mathcal{R}} P_{ijr}(\nu_i^b; \rho)^{n_{ir}} \right]. \tag{18}$$

Here, j denotes the household's washing machine, and ν_i^b is a draw of its temperature and duration preferences. The multinomial coefficient can be omitted because it does not depend on the parameters. Integrating over tastes at the household level preserves preference heterogeneity across reported washes. I then compute household-specific inclusive values for product demand.

Product demand is estimated by aggregate-share GMM using the inversion of Berry et al. (1995) and the implementation of Conlon and Gortmaker (2020). With demand residual $\xi(\theta)$ and instruments Z, the criterion is

$$g(\theta) = \frac{1}{n} Z'\xi(\theta), \qquad \widehat{\theta} = \arg\min_{\theta} g(\theta)' W g(\theta). \tag{19}$$

I use a one-step weighting matrix and aggregate moments, without additional micro moments. Income, stain intensity, and temperature and duration tastes are integrated using 625 weighted agents per market. The product-specific conditional-efficiency component enters demand directly but is not used to construct additional characteristic-sum instruments.

Given the demand estimates and product ownership, the multiproduct Bertrand first-order conditions recover markups and marginal costs. I regress log recovered marginal cost on the variables in equation (13). Its residual and the demand residual determine the joint shock distribution used to integrate design-stage profits. Operating profits, design costs, and the monetary equivalents of entry costs are expressed in euros, consistently with the price normalization in demand.

Instruments. The instruments follow the timing of the model. Product characteristics and program designs are chosen before the product-market demand shock ξ_{jmt} and marginal-cost shock ω_{jt} are realized, whereas prices are set after both shocks are observed. Price is therefore endogenous. I assume that predetermined product and program characteristics, including conditional efficiency, are orthogonal to the contemporaneous demand innovation after conditioning on the included characteristics and fixed effects. This orthogonality is a maintained identifying assumption rather than a consequence of timing alone.

I instrument for price using standard BLP instruments constructed from sums of capacity, spin speed, and displayed electricity consumption among products offered by the same firm and by rival firms. Because price sensitivity varies with household income, I also interact these instruments with mean household income in market (m, t). These interactions provide additional variation to help identify income heterogeneity in price sensitivity.

Because the inclusive value $\mathcal{I}_{ij}$ varies across households, β^I enters the nonlinear component of aggregate product demand. I use the market-average inclusive value

$$Z^I_{jmt} = \int \mathcal{I}_{ij}\, dF_{mt}(i) \tag{20}$$

as an additional instrument. It summarizes variation in expected program-use value across products and markets. Its validity rests on the assumption that program designs and the household distribution are orthogonal to the contemporaneous demand innovation.

Collecting the excluded instruments and exogenous product characteristics in Z_{jmt}, the aggregate moments satisfy

$$\mathbb{E}\left[Z_{jmt}\xi_{jmt}(\theta)\right] = 0. \tag{21}$$

Identification relies on the exogeneity and sufficient variation of these instruments, following the general framework for differentiated-products demand in Berry and Haile (2014).

6.2 Supply Estimation

I estimate supply parameters in two steps. First, I estimate the design costs governing program-level input choices conditional on the observed product portfolio. Second, taking the estimated design costs as given, I estimate the entry costs governing energy-class choices within each product line.

Design costs. I estimate the parameters (κ^d, C^d) in equation (14) for four program-input groups: Cotton 40°C duration, Cotton 60°C duration, Eco duration, and Eco temperature. Each group has its own linear coefficient, while Cotton 40°C and Eco duration share a quadratic coefficient, giving seven free parameters. This pooling restriction addresses limited identification of Eco-duration curvature due to bunching at the regulatory duration cap. Product-specific conditional efficiency remains fixed. I use a minimum distance (MD) estimator that minimizes a weighted distance between observed targets and their model-implied counterparts, following the moment-matching approach in Hall and Rust (2003), Goettler and Gordon (2011), Chen and Xu (2022), and Su et al. (2026).[24] For each candidate parameter vector, I solve the program-design game by iterating firms' constrained best responses, using sequential least squares programming (SLSQP) for each firm's optimization problem. Appendix F.2 describes the reference designs and estimation sample, and Appendix F.3 details the solution algorithm.

The MD estimator solves

$$(\widehat{\kappa}^d, \widehat{C}^d) = \arg \min_{\substack{\kappa^d \in \mathbb{R}^4,\, C^d \in \mathbb{R}^4_{++} \\ C^d_{40} = C^d_d}} \sum_{n=1}^{N} w_n \left[\widehat{m}_n(\kappa^d, C^d) - m_n \right]^2. \tag{22}$$

Here, N is the number of moment conditions, m_n denotes an observed target or a zero-valued first-order-condition restriction, and $\widehat{m}_n(\kappa^d, C^d)$ is its model-implied counterpart. Fixed weights standardize input discrepancies using empirical input standard deviations and marginal-profit discrepancies using fixed marginal-profit scales. Each input group receives equal total weight. Within each group, one-third of the weight is assigned to baseline restrictions and two-thirds to design moments. For duration groups, baseline weight is split equally between marginal-profit and mean-duration restrictions. Design weight is split equally between class means and individual input changes, with equal weight within each block. I use three sets of moments:

- *Baseline first-order-condition and duration restrictions.* I match average marginal operating profits at the A^{+++} reference designs to the linear cost coefficients. For input group g, the

[24] An alternative is to recover design costs from firms' first-order conditions (Fan, 2013; Barwick et al., 2024; Ito et al., 2026). In my setting, binding regulatory constraints introduce product-specific shadow values. Solving the constrained design problem incorporates these shadow values implicitly, avoiding the need to estimate them separately.

first-order-condition discrepancy is

$$\kappa_g^d - \frac{1}{|\mathcal{J}_g^0|} \sum_{j \in \mathcal{J}_g^0} \frac{\partial \Pi_{f(j)}^{\text{op}}(q^{\text{ref}})}{\partial q_{jg}},$$

where $\Pi_{f(j)}^{\text{op}}$ denotes operating profit before design costs and $\mathcal{J}_g^0$ contains eligible baseline observations. Marginal profits are calculated from the estimated demand and pricing model. For the three duration groups, I also match observed and optimized mean durations among eligible A^{+++} products.

- *Input changes at first class introduction.* I match observed input changes, $q_{jg} - q_{jg}^{\text{ref}}$, to predicted changes at the first introduction of each energy class on a product line. Repeated introductions inherit existing designs and do not provide separate adjustment targets.

- *Class-specific mean inputs.* I match observed mean inputs,

$$\frac{1}{|\mathcal{J}_{g,c}|} \sum_{j \in \mathcal{J}_{g,c}} q_{jg},$$

to their model-predicted counterparts, where $\mathcal{J}_{g,c}$ contains eligible observations for input group g and energy class c.

The parameters are jointly estimated from the full set of moments, although different moment sets provide the main identifying information for different parameters. The baseline first-order-condition restrictions primarily inform the linear coefficients κ_g^d. I treat the matched A^{+++} designs as optimal under the pre-reform environment represented by 2020Q1 and use their own input levels as references for these restrictions. At the reference, the quadratic cost term has zero derivative, leaving κ_g^d as the marginal design cost. The EEI distribution in Figure A7 supports treating the old-label EEI constraint as slack for these products. For interior choices without other binding constraints, firms' first-order conditions therefore link the linear cost coefficients to marginal operating profits.

The input-change moments primarily inform the curvature parameters C_g^d, which govern how marginal design costs change with deviations from the reference. Conditional on the linear coefficients and other determinants of input choices, larger observed adjustments generally favor lower curvature. This relationship is less informative for Eco duration because many products must shorten their pre-reform cycles to satisfy the cap. Conditional on entry, a range of sufficiently large adjustment costs can lead firms to make the smallest required reduction, producing the same bunching at the cap. I therefore impose $C_d^d = C_{40}^d$, allowing Cotton 40°C duration adjustments to inform the curvature applied to Eco duration.

The class-specific mean-input moments place additional weight on systematic differences between observed and predicted designs across energy classes. They complement the baseline and input-

change moments in the joint estimation rather than primarily identifying a particular cost parameter. Mean energy consumption is reported as a model-fit diagnostic but does not enter the estimation criterion.

Entry costs. I estimate the common launch cost κ^e, six switching premiums $\{\kappa^{e,\text{sw}}_{g,h}\}_{g\in\{A,B,CD\},\, h\in\{\text{tran,post}\}}$, and the profit coefficient β_π, collecting these eight parameters in θ. Switching premiums are pooled across classes C and D within each period, as specified in equation (17). Variation in product introductions, class-switching histories, and expected profitability disciplines these parameters. Estimation requires equilibrium beliefs about rivals' portfolios, which determine the expected payoff from each introduction portfolio.

Under the incomplete-information framework of Seim (2006), the maintained portfolio-level shock assumption yields logit portfolio probabilities. Define the deterministic expected payoff as

$$\overline{W}_{ft}(a_{ft};\theta,\mathcal{P}_{-f,t}) = \beta_\pi \mathbb{E}_{\mathcal{P}_{-f,t}}\left[W_{ft}(a_{ft},a_{-f,t})\right] - C^{\text{entry}}_{ft}(a_{ft};\theta), \tag{23}$$

where W_{ft} is obtained from the program-design problem in equation (15). The portfolio choice probability is

$$\mathcal{P}_{ft}(a_{ft};\theta,\mathcal{P}_{-f,t}) = \frac{\exp\{\overline{W}_{ft}(a_{ft};\theta,\mathcal{P}_{-f,t})\}}{\sum_{a'_{ft}\in\mathcal{F}_{ft}}\exp\{\overline{W}_{ft}(a'_{ft};\theta,\mathcal{P}_{-f,t})\}}. \tag{24}$$

Direct estimation faces two computational challenges. First, a firm with $|\mathcal{L}_{ft}|$ potential product lines has $5^{|\mathcal{L}_{ft}|}$ feasible portfolios, since each line permits no entry or entry in class A, B, C, or D. Computing the portfolio normalizer therefore becomes costly even when rivals' portfolios can be simulated. Second, evaluating each candidate market portfolio requires solving the constrained program-design game.

I address the first challenge by conditioning on the firm's choices on its other product lines. For line $\ell \in \mathcal{L}_{ft}$,

$$\mathcal{P}_{f\ell t}\left(a \mid a_{f,-\ell,t};\theta,\mathcal{P}_{-f,t}\right) = \frac{\exp\left\{\overline{W}_{ft}\left((a,a_{f,-\ell,t});\theta,\mathcal{P}_{-f,t}\right)\right\}}{\sum_{a'\in\mathcal{A}}\exp\left\{\overline{W}_{ft}\left((a',a_{f,-\ell,t});\theta,\mathcal{P}_{-f,t}\right)\right\}}, \tag{25}$$

where $a_{f,-\ell,t}$ denotes the firm's choices on all other lines and $\mathcal{A} = \{0,A,B,C,D\}$. Conditioning cancels the full portfolio normalizer, leaving five alternatives while preserving interactions among products. Given rival beliefs, I maximize the conditional composite log likelihood:

$$\mathcal{L}(\theta;\mathcal{P}) = \sum_{(f,\ell,t)\in\mathcal{S}^{\text{entry}}} \log \mathcal{P}_{f\ell t}\left(a^{\text{obs}}_{f\ell t} \mid a^{\text{obs}}_{f,-\ell,t};\theta,\mathcal{P}_{-f,t}\right), \tag{26}$$

where $\mathcal{S}^{\text{entry}}$ contains the eligible line-quarter entry opportunities.

Expected profits still depend on rivals' full portfolio distributions. I simulate these distributions using Gibbs sampling, sequentially updating each line's entry and energy-class choice conditional on the firm's current choices on its other lines (Casella and George, 1992). Parameters and rival beliefs remain fixed during each sampling step. I then alternate between estimating θ given rival beliefs and updating those beliefs using simulated portfolios until parameters and monitored conditional probabilities stabilize (Aguirregabiria and Mira, 2007).

The theoretical results in Appendix G establish that, under compatibility and positivity conditions, the conditional representation preserves the equilibrium set and full and conditional likelihood have the same population identified set within a common equilibrium family. Monte Carlo experiments in Appendix G.2 show that the two approaches produce nearly identical parameter estimates. Appendix G.3 provides sample-selection and implementation details.

To address the second computational challenge, I train a neural-network surrogate on numerical solutions of the program-design game. The surrogate predicts optimal program inputs and firms' operating profits net of design costs across market portfolios, avoiding repeated constrained optimization during entry estimation. Appendix F.1 describes the network and its predictive fit.

7 Estimation Results

7.1 Demand Estimates

Panel A of Table 3 reports estimates of the program-use utility function in equation (6). On average, households value higher temperatures, longer cycles, and greater water use, holding other program attributes fixed. Households with a larger share of heavily stained laundry place less value on temperature and more value on duration. Duration preferences vary substantially across households, with an estimated standard deviation of 2.248 per hour. This heterogeneity shapes substitution across programs: holding other attributes fixed, shortening Eco by 30 minutes reduces its predicted usage share by 1.65 percentage points and raises the combined Cotton share by 1.61 percentage points.[25] This response is consistent with households valuing the cleaning performance associated with longer cycles, despite the additional time required.

Panel B of Table 3 reports estimates of the product-choice utility function in equation (9). Price sensitivity decreases with income: the log-income coefficient is -0.578 and significant at the 5 percent level. The coefficient on product-specific conditional energy efficiency, z_j, is -1.147 and significant at the 1 percent level. Since lower values of z_j indicate lower electricity consumption conditional on program inputs, this estimate associates greater conditional efficiency with higher

[25]The simulation uses the 117 households in the new-label survey, weighted by reported wash counts, and holds all other program attributes fixed.

demand, holding the included characteristics, program value, and energy class fixed.

The estimates also imply a substantial purchase premium for class A. Relative to class D, the utility coefficients are 0.956 for class A, 0.186 for B, and 0.172 for C, all significant at the 1 percent level. The class-A coefficient implies a willingness to pay of approximately €145 relative to class D for a household at the sample mean log income, holding other characteristics fixed. This valuation is higher for households with greater income because they are less price sensitive. Finally, the coefficient on the expected value of the wash-program menu is 0.147 and significant at the 1 percent level, indicating that households are more likely to purchase machines offering programs better suited to their preferences.

7.2 Supply Estimates

Panel A of Table 4 reports the marginal-cost estimates from equation (13). Greater conditional energy efficiency is associated with higher production costs. The coefficient $\gamma_z = -0.116$ implies that a reduction in z_j corresponding to 10 percent lower electricity consumption at fixed program inputs is associated with approximately 1.23 percent higher predicted marginal cost. Relative to 10 kg machines, predicted marginal costs are lower for 7, 8, and 9 kg machines. The coefficient on 1400 rpm is negative, while that on 1600 rpm is close to zero, relative to the omitted 1200 rpm category. The regression includes firm and quarter fixed effects.

Panel B reports the design-cost coefficients. For input group g, a deviation h from its reference incurs a cost of $\kappa_g^d h + C_g^d h^2$. A ten-minute extension from the reference costs approximately € 843 for Cotton 40°C, € 2,870 for Cotton 60°C, and € 580 for each Eco-duration coordinate. Extending both full- and half-load Eco cycles by ten minutes therefore costs approximately € 1,159. These are costs per introduction, excluding discrete entry costs and changes in operating profits. They measure departures from the reference rather than incremental adjustments to an already modified design.

I impose a common quadratic coefficient for Cotton 40°C and Eco duration, $C_{40}^d = C_d^d = 5.91$, because bunching at the Eco-duration cap limits separate identification of its curvature. Their linear coefficients differ: Eco's is close to zero, whereas those for both Cotton programs are positive. Cotton 60°C has a larger quadratic coefficient, $C_{60}^d = 23.10$, implying that its marginal adjustment cost rises more rapidly with departures from the reference. The Eco-temperature quadratic coefficient is 50.51, although its magnitude is not directly comparable with duration coefficients because the inputs have different units. Equilibrium adjustments also depend on regulatory constraints, household preferences, and the other program inputs.

Panel C reports a common launch cost of $\kappa^e = 2.118$ and a scaled profit coefficient of $10^6\beta_\pi = 0.478$. The positive profit coefficient indicates that greater expected operating profits net of design

Table 3: Program-Use and Product-Demand Estimates

Variable	Parameter	Estimate	Std. Error
Panel A: Program-use estimates			
Program attributes			
Temperature	ρ_0^T	0.046	(0.004)
Duration (hours)	ρ_0^d	0.582	(0.064)
Water consumption	ρ_w	0.016	(0.001)
Representative load (kg)	ρ_L	0.373	(0.027)
Household preference variation			
Temperature × high-stain share	ρ_1^T	−0.041	(0.012)
Duration × high-stain share	ρ_1^d	0.362	(0.171)
Temperature taste SD	σ_T	0.064	(0.002)
Duration taste SD	σ_d	2.248	(0.068)
Panel B: Product-demand estimates			
Linear demand coefficients			
Conditional energy efficiency	β_z^x	−1.147	(0.227)
Capacity: 7 kg	$\beta_{7\text{kg}}^x$	0.309	(0.251)
Capacity: 8 kg	$\beta_{8\text{kg}}^x$	0.220	(0.185)
Capacity: 9 kg	$\beta_{9\text{kg}}^x$	0.412	(0.102)
Spin speed: 1400 rpm	β_{1400}^x	0.140	(0.125)
Spin speed: 1600 rpm	β_{1600}^x	0.392	(0.126)
Energy class A $\times \mathbf{1}_{\text{new},t}$	β_A	0.956	(0.131)
Energy class B $\times \mathbf{1}_{\text{new},t}$	β_B	0.186	(0.064)
Energy class C $\times \mathbf{1}_{\text{new},t}$	β_C	0.172	(0.049)
Household-specific coefficients			
Log price-sensitivity intercept	$\bar{\alpha}$	1.619	(1.057)
Log-income effect on price sensitivity	α^y	−0.578	(0.278)
Program value	β^I	0.147	(0.045)

Notes: Panel A pools the old- and new-label surveys and reports Hessian-based standard errors in parentheses. The outside option is normalized to zero. Temperature and duration coefficients satisfy $\rho_i^q = \rho_0^q + \rho_1^q \tau_i + \sigma_q \nu_i^q$ for $q \in \{T, d\}$. Panel B reports robust standard errors. Prices and income are measured in hundreds of euros, and price sensitivity is $\alpha_i = \exp(\bar{\alpha} + \alpha^y \log y_i)$, entering utility as $-\alpha_i p_{jmt}$. Program value is the household-specific inclusive value $\mathcal{I}_{ij}$. Conditional energy efficiency is the product component z_j from the energy equation. Class D is omitted; label effects apply from 2021Q1. Demand covers 2020Q1–2023Q4.

costs increase the probability of introduction, although the coefficient is imprecisely estimated. Entry costs are expressed in normalized payoff units; dividing them by β_π gives their monetary equivalents.

During the transition period, switching premiums are 2.576 for class A, 3.434 for class B, and 1.609 for classes C and D. After the reform, the corresponding premiums are 1.892, 3.494, and 4.735. These premiums are additional to the common launch cost and apply when the introduced class was absent from the line's most recent introduction quarter, favoring persistence in recently

introduced classes. The lower post-reform premium for class A may reflect improved components or learning that makes highly efficient models cheaper to configure, although the estimates do not separately identify these mechanisms.

These parameters capture effective introduction costs, including certification, marketing, and the opportunity cost of scarce engineering resources. Many lines with positive expected operating profits do not introduce products, consistent with firms facing costs of hiring additional engineers or diverting them from other projects. Because a modeled product may encompass several commercial variants, these costs can cover multiple models. They may also absorb omitted continuation values and errors in predicted profits.

Table 4: Supply Estimates

Variable	Parameter	Estimate	Std. Error
Panel A: Marginal costs			
Conditional energy efficiency	γ_z	-0.116	(0.047)
Capacity: 7 kg	$\gamma_{7\text{kg}}$	-0.552	(0.016)
Capacity: 8 kg	$\gamma_{8\text{kg}}$	-0.347	(0.015)
Capacity: 9 kg	$\gamma_{9\text{kg}}$	-0.146	(0.015)
Spin speed: 1400 rpm	γ_{1400}	-0.179	(0.028)
Spin speed: 1600 rpm	γ_{1600}	-0.009	(0.029)
Panel B: Design costs			
Linear coefficients			
Cotton 40°C duration	κ^d_{40}	25.19	(9.23)
Cotton 60°C duration	κ^d_{60}	56.03	(22.83)
Eco duration	κ^d_d	-1.13	(5.13)
Eco temperature	κ^d_T	40.82	(13.89)
Quadratic coefficients			
Cotton 40°C and Eco duration	$C^d_{40} = C^d_d$	5.91	(1.82)
Cotton 60°C duration	C^d_{60}	23.10	(8.12)
Eco temperature	C^d_T	50.51	(20.79)
Panel C: Entry costs			
Common launch cost	κ^e	2.118	(1.259)
Profit coefficient (scaled)	$10^6\beta_\pi$	0.478	(0.291)
Transition switching premiums			
Class A	$\kappa^{e,\text{sw}}_{A,\text{tran}}$	2.576	(1.574)
Class B	$\kappa^{e,\text{sw}}_{B,\text{tran}}$	3.434	(2.124)
Classes C and D	$\kappa^{e,\text{sw}}_{CD,\text{tran}}$	1.609	(0.959)
Post-reform switching premiums			
Class A	$\kappa^{e,\text{sw}}_{A,\text{post}}$	1.892	(0.768)
Class B	$\kappa^{e,\text{sw}}_{B,\text{post}}$	3.494	(1.048)
Classes C and D	$\kappa^{e,\text{sw}}_{CD,\text{post}}$	4.735	(1.602)

Notes: Panels A and C use 7,496 product observations and 469 line-quarter observations, respectively. Design costs are in euros; entry costs are in normalized payoff units, with the profit coefficient reported as $10^6\beta_\pi$. Standard errors use HC1 for marginal costs, a firm-clustered minimum-distance sandwich for design costs, and a conditional composite-likelihood HC1 sandwich for entry costs.

7.3 Model Fit

I assess design fit using the equilibrium program inputs obtained from the same constrained solver used in minimum-distance estimation, conditional on the observed product portfolio and reported design-cost coefficients. Figure G6 and Table A6 in Appendix G.4 compare these predictions with the data. The model reproduces the main patterns in endogenous program design: shorter Eco cycles and generally longer Cotton cycles among new-label products, together with lower Eco temperatures for higher-rated machines. Fitted class means generally track their observed counterparts, indicating that the common design-cost specification captures systematic differences across programs and energy classes. The model also captures the broad patterns in energy consumption, which does not enter the design-cost estimation criterion and provides an additional check on the fitted input combinations.

For entry, I simulate rival portfolios using Gibbs sampling at the estimated parameters and equilibrium beliefs, then calculate each line's conditional choice probabilities given the firm's observed choices on its other lines. Averaging these probabilities gives predicted entry rates and energy-class shares. Figure G7 in Appendix G.4 shows that the model captures the prevalence of no entry, the prominence of class A among introductions, and persistence in previously introduced classes. Predicted quarterly entry rates also follow the broad movements in observed introductions, although they are smoother than realized outcomes. Together, these comparisons indicate that the model captures the main patterns in firms' program-design and introduction decisions.

8 Counterfactuals

I use the estimated model to evaluate whether alternative policies can reduce electricity consumption and increase consumer welfare. I consider five counterfactuals: no 2021 reform, removing the Eco-duration cap, removing labels and efficiency restrictions, broadening the efficiency index to include Cotton programs, and introducing an energy-based producer tax. Table 5 summarizes their specifications.

Preferences and estimated cost parameters remain fixed. Firms and households adjust their choices conditional on observed quarterly histories, except that the no-reform counterfactual fixes the pre-reform portfolio and designs while allowing prices and household choices to adjust. I aggregate outcomes over 2020Q1–2023Q4 and compare them with model-implied current policy. Appendix H provides implementation details.

Table 5: Counterfactual Policy Experiments

Counterfactual	Policy specification
No duration cap	Removes the Eco-duration cap while retaining energy labels and efficiency requirements.
No labels or efficiency restrictions	Removes label utility and efficiency-index restrictions while retaining the Eco-duration cap.
No 2021 reform	Retains the pre-reform labeling and regulatory environment.
Weighted EEI	Broadens the efficiency index to include Cotton programs, retains labels and class thresholds, and removes the duration cap.
Energy-based producer tax	Replaces labels, efficiency restrictions, and the duration cap with a producer tax based on Eco and Cotton energy consumption. Revenue is rebated to households.

Notes: All counterfactuals are compared with model-implied current policy. The no-reform counterfactual fixes the pre-reform portfolio and designs while allowing prices and household choices to adjust. The remaining counterfactuals also allow product introductions and program designs to adjust conditional on observed quarterly histories. Preferences and estimated cost parameters remain fixed.

8.1 Welfare Accounting

Energy-label utility may partly reflect perceived electricity savings. I assume that households interpret the Eco-based label as information about the energy efficiency of the whole machine. To avoid counting perceived savings alongside actual electricity savings, I project the estimated class coefficients onto class-level Eco energy savings relative to class D:

$$\widehat{\beta}_c = \widehat{\gamma} x_c + r_c, \qquad x_c = \bar{e}^E_D - \bar{e}^E_c, \qquad c \in \{A, B, C\}. \tag{27}$$

Energy is measured per loaded kilogram using common capacity and quarter weights, with $x_D = 0$. I interpret the fitted component as perceived energy-saving utility relative to class D and retain the residual as nonpecuniary label value. This interpretation is an accounting assumption. The correction uses Eco energy because it underlies the label, whereas actual electricity expenditure depends on the energy consumed by each program and how frequently households use it.

For a given equilibrium portfolio, consumer welfare is

$$\begin{aligned} \mathcal{CW}^{\text{net},s}_{mt} = M_{mt} \int \frac{1}{\alpha_i} \left[\log\left(1 + \sum_{j \in \mathcal{J}^s_{mt}} \exp\{\bar{u}^s_{ijmt}\} \right) - \sum_{j \in \mathcal{J}^s_{mt}} P^s_{ijmt} \lambda^s_{jmt} \widehat{\gamma} x_{c(j)} \right] dF_{mt}(i) \\ + \mathcal{R}^s_{mt} - \mathcal{C}^s_{E,mt}, \end{aligned} \tag{28}$$

where α_i is the marginal utility of one euro, P^s_{ijmt} is the purchase probability, λ^s_{jmt} indicates whether new-label utility is active, $\mathcal{R}^s_{mt}$ is the tax rebate, and $\mathcal{C}^s_{E,mt}$ is discounted electricity expenditure. The correction removes the perceived savings associated with the label, not an observed Eco electricity bill. It uses the original purchase probabilities and does not change household choices. The program inclusive value remains in purchase utility, so no separate washing-benefit term is added. This accounting treats its estimated purchase-stage valuation as welfare relevant and does not separately correct for possible inattention to future program use.

Electricity expenditure covers purchased machines and machines retained by nonbuyers. Buyer electricity depends on purchases, program choices, and the electricity consumed per cycle by each covered program. Nonbuyers use a common reference basket based on the sales-weighted 2021Q1 portfolio. I value electricity at € 0.30 per kWh over a 12-year service life with a 5 percent discount rate. These assumptions and washing frequency are common across policies. Appendix I provides the full accounting and sensitivity analysis.

Operating profits equal revenue less variable production costs and producer-tax payments, excluding entry and design costs. Tax revenue is returned once to households through lump-sum transfers.

8.2 Policy Effects on Product Design and Welfare

Figures 4 and 5 show how the policies affect product design, electricity expenditure, consumer welfare, and operating profits. Table A7 reports the corresponding levels and changes. The central mechanism is that households value program attributes as well as energy efficiency. Regulating Eco therefore changes firms' incentives to design the entire program menu, while actual electricity use depends on which programs households choose. Attribute statistics describe first firm–line–class introductions, except in the no-reform benchmark, which uses the fixed pre-reform portfolio.

Removing the duration cap. The duration cap restricts one way firms can combine valued program attributes with energy efficiency. Without the cap, firms offer Eco cycles that are 51.93 minutes longer and 1.81°C cooler. Longer Eco cycles can better serve households that value additional washing time. The accompanying decline in Cotton durations is consistent with firms having less incentive to accommodate these preferences through Cotton programs: Cotton 40°C and Cotton 60°C durations decline by 3.0 and 2.4 percent.

Longer Eco duration alone raises energy use in the estimated engineering relationship, while lower temperature works in the opposite direction. The net outcome also reflects changes in the composition of introductions. Electricity consumption per cycle falls by 3.0 percent for Eco, 2.1 percent for Cotton 40°C, and 1.7 percent for Cotton 60°C. The shorter Cotton cycles and lower

energy consumption are consistent with relaxing the cap reducing incentives to shift energy-intensive program attributes toward Cotton.

The displayed aggregate comparison shows 1.11 percent lower electricity expenditure and a consumer-welfare gain of € 63.24 million. Operating profits rise by 6.51 percent. Retaining labels preserves the commercial reward for efficiency, while relaxing the duration restriction gives firms more flexibility to offer program menus that households value.

Removing labels and efficiency restrictions. Removing labels and efficiency requirements weakens the regulatory and commercial rewards for concentrating efficiency improvements in Eco. Eco temperature rises by 14.2 percent, and its electricity consumption per cycle increases by 16.2 percent. Cotton 40°C and Cotton 60°C consumption instead falls by 6.0 and 1.8 percent. The model therefore predicts a reallocation of energy use across programs when firms no longer compete for energy classes determined by Eco performance.

This reallocation does not deliver an aggregate electricity saving: expenditure rises by 0.51 percent. Removing the regime can reduce incentives to differentiate Eco from Cotton, but also weakens incentives for Eco efficiency. This trade-off helps explain why eliminating the regime need not achieve the same outcome as selectively relaxing its duration cap.

The displayed consumer-welfare gain is € 59.52 million, while operating profits fall by 13.10 percent. Household benefits from changes in prices and program menus can coexist with higher electricity expenditure. The welfare accounting treats residual label value as welfare relevant wherever labels remain, so the comparison does not assume that all label utility lacks welfare value.

No 2021 reform. The preceding experiments examine the duration cap and the labeling and efficiency regime separately. The no-reform counterfactual provides a broader benchmark by retaining the pre-reform portfolio and designs while allowing prices and household choices to adjust. Eco electricity consumption per cycle is 23.6 percent higher than under current policy, but Cotton 60°C consumption is 8.4 percent lower, and aggregate electricity expenditure is 3.02 percent lower. This contrast reinforces the importance of evaluating efficiency across the programs households actually use. The displayed consumer-welfare gain is € 231.13 million, while operating profits are 12.39 percent lower. Because the portfolio and designs are fixed, these results describe retaining the pre-reform product offering rather than the industry's unrestricted response to the absence of the reform.

Broadening the efficiency index. Broadening the index addresses the mismatch between the program used for regulation and the programs households use. Under the Eco-only index, reducing

Cotton energy consumption does not help a machine obtain a better label. Including Cotton creates that incentive: firms can improve their regulatory performance by reducing energy use across the menu. The experiment uses $0.80e_{jE} + 0.10e_{j40} + 0.10e_{j60}$, retains class thresholds, and removes the duration cap. It therefore changes both regulatory coverage and effective stringency.

The strongest adjustment occurs in Cotton 40°C, whose duration falls by 31.9 percent and electricity consumption by 25.0 percent. Eco becomes longer and cooler, with electricity consumption falling by 17.8 percent; Cotton 60°C consumption falls by 3.5 percent. These responses are consistent with broader regulatory coverage redirecting firms' design incentives toward programs outside the original test.

Aggregate electricity expenditure falls by 0.33 percent, and the displayed consumer-welfare gain is € 76.59 million. Operating profits rise by 5.05 percent. The aggregate electricity response is smaller than the improvements in entrant attributes because it also depends on sales, program use, and retained machines. In particular, shorter Cotton cycles can change their attractiveness to households that value duration. Broader coverage improves energy performance across the program menu, but its effect on actual use also depends on these household responses. The specified weights are not established as optimal.

An energy-based producer tax. The producer tax gives firms a continuous incentive to reduce assessed energy consumption rather than rewarding compliance with discrete class thresholds. It replaces labels, efficiency restrictions, and the duration cap with

$$\mathcal{T}_j = \frac{\tau_E e_{jE} + \tau_C(e_{j40} + e_{j60})}{3}, \qquad \tau_E = \tau_C = 50. \tag{29}$$

The levy is charged per machine sold, and revenue is rebated to households. Household electricity prices remain unchanged.

The tax encourages savings across the assessed programs, but removing labels and efficiency requirements simultaneously weakens the previous incentive for Eco efficiency. Eco electricity consumption per cycle rises by 4.3 percent, while Cotton 40°C consumption falls by 12.0 percent and Cotton 60°C consumption changes little. These outcomes reflect both the tax incentive and the removal of the existing regulatory regime.

The levy targets average program energy consumption rather than the electricity households ultimately use. Firms respond to the tax weights, whereas actual consumption depends on household program choices and purchases. Lower assessed energy intensity therefore need not translate into lower electricity expenditure. In this experiment, expenditure rises by 1.36 percent, consumer welfare falls by € 74.75 million despite rebates, and operating profits decline by 45.13 percent. Rebates return tax payments to households but do not undo changes in prices, product availability,

or program design.

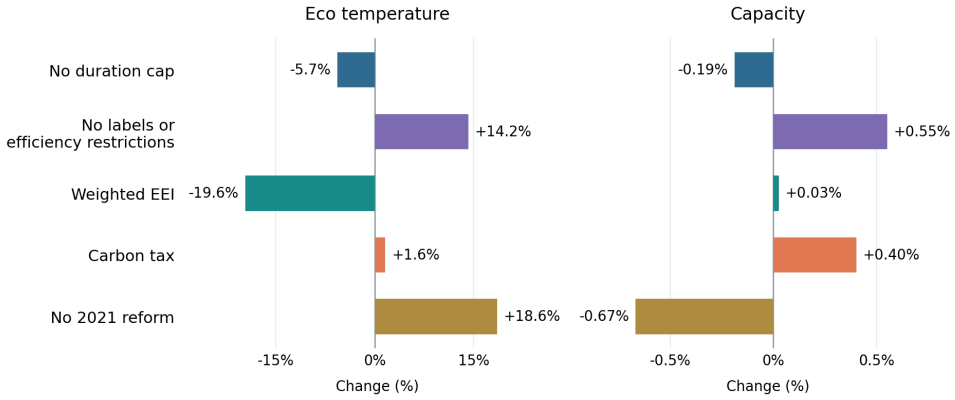
Among the tax rates evaluated under the same deregulated environment, zero tax yields the highest consumer welfare net of electricity expenditure. This finding concerns the specified product tax and evaluated grid; a tax more closely aligned with realized electricity use could produce different results.

9 Conclusion

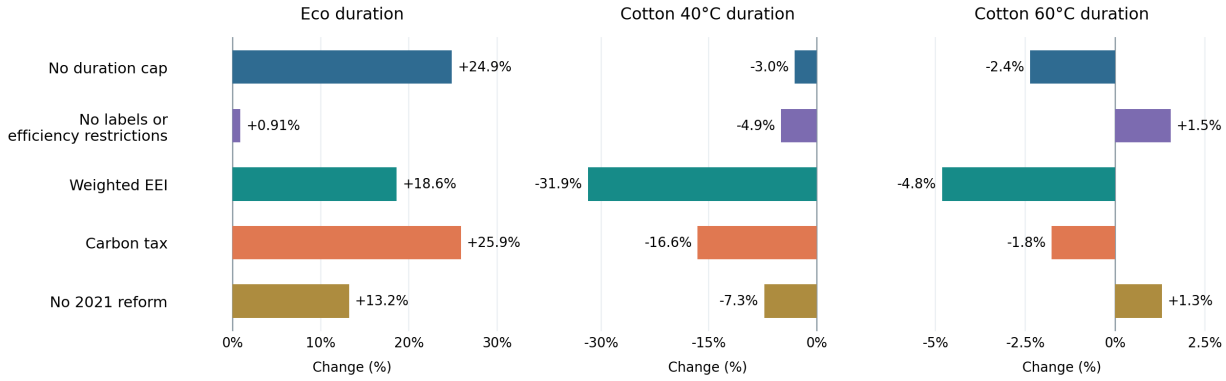
This paper studies how energy labels and ecodesign constraints affect product design when regulatory tests capture only one of several operating modes. I examine the 2021 EU washing-machine reform in Germany, which rescaled energy labels and imposed an Eco-duration cap. The evidence shows that energy consumption declined in the regulated Eco program but rose in hotter Cotton programs. Because Eco accounts for only a small share of household wash cycles, improvements in the tested program need not translate into comparable savings in actual use.

The estimated model links these responses to firms' design incentives and households' program choices. Relaxing the Eco-duration cap allows firms to offer longer Eco cycles and reduces incentives to accommodate preferences for longer cycles through Cotton programs. Broadening the efficiency index directly rewards energy savings outside Eco. Both policy experiments lower electricity expenditure and improve consumer welfare under the evaluated specifications. By contrast, the energy-based producer tax raises electricity expenditure and lowers consumer welfare despite revenue rebates. These findings highlight the importance of regulatory flexibility and the alignment between assessed efficiency and household use: effective policy must account for firms' design responses across the full program menu and consumers' choices among those programs.

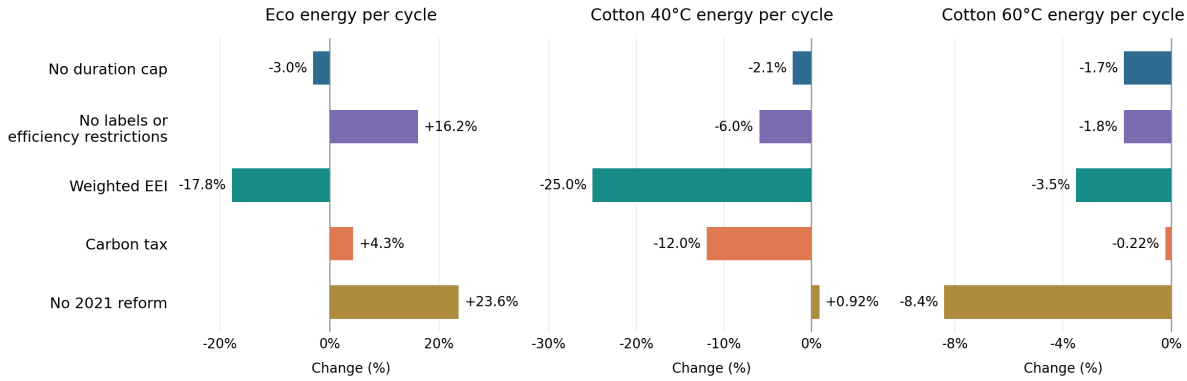
Figure 4: Product Attributes and Energy Consumption under Counterfactual Policies



(a) Eco temperature and capacity



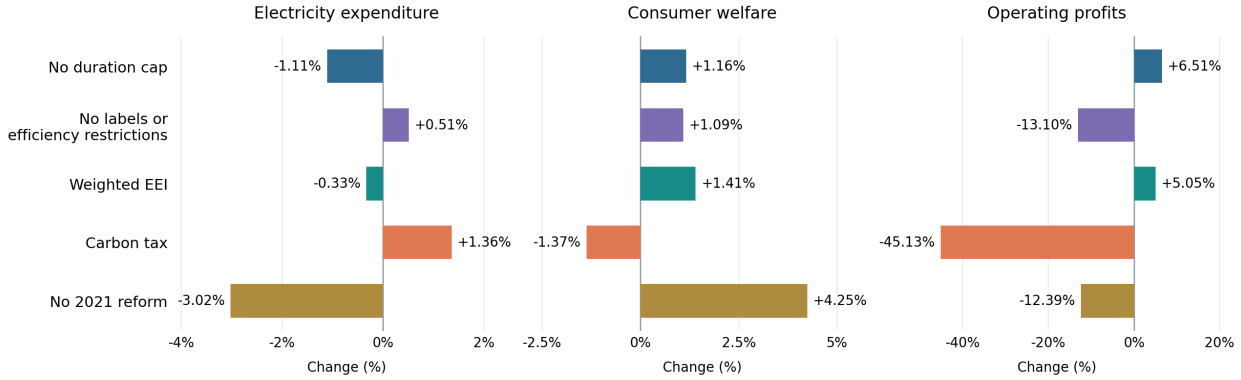
(b) Program duration



(c) Energy consumption per cycle

Notes: Percentage changes relative to the model-implied current-policy baseline. Attributes are unweighted means of first class introductions within product lines. Eco temperature and duration refer to full load; Eco energy consumption is load weighted. Horizontal scales differ across outcomes.

Figure 5: Welfare, Electricity Expenditure, and Profits under Counterfactual Policies



Notes: Changes are relative to the model-implied current-policy baseline, expressed as percentages of each outcome's absolute baseline value. Consumer welfare is net of electricity expenditure and includes tax rebates. Electricity expenditure covers buyers and retained machines. Horizontal scales differ across outcomes.

References

Aguirregabiria, V. and Mira, P. (2007). Sequential estimation of dynamic discrete games. *Econometrica*, 75(1):1–53.

Alborzi, F., Boyano, A., Cordella, M., Espinosa, N., Graulich, K., Hook, I., Rüdenauer, I., Stamminger, R., and Villanueva, A. (2017). Ecodesign and energy label for household washing machines and washer dryers: Preparatory study – final report. Technical Report EUR 28809 EN, Publications Office of the European Union, Luxembourg. JRC108604.

Baker, G. P. (1992). Incentive contracts and performance measurement. *Journal of political Economy*, 100(3):598–614.

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

Barrett, S. R. H., Speth, R. L., Eastham, S. D., Dedoussi, I. C., Ashok, A., Malina, R., and Keith, D. W. (2015). Impact of the volkswagen emissions control defeat device on us public health. *Environmental Research Letters*, 10(11):114005.

Barwick, P. J., Kwon, H.-S., and Li, S. (2024). Attribute-based subsidies and market power: an application to electric vehicles. Technical report, National Bureau of Economic Research.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Berry, S. T. and Haile, P. A. (2014). Identification in differentiated products markets using market level data. *Econometrica*, 82(5):1749–1797.

Casella, G. and George, E. I. (1992). Explaining the gibbs sampler. *The American Statistician*, 46(3):167–174.

Chen, Y. and Xu, D. Y. (2022). A structural empirical model of R&D investment, firm heterogeneity, and industry evolution. Working Paper 29733, National Bureau of Economic Research.

Conlon, C. and Gortmaker, J. (2020). Best practices for differentiated products demand estimation with pyblp. *The RAND Journal of Economics*, 51(4):1108–1161.

Czyzewski, G. (2022). Energy savings 2021: New european energy label and ecodesign requirements for washing machines. In Bertoldi, P., editor, *Energy Efficiency in Domestic Appliances and Lighting: Proceedings of the 10th International Conference (EEDAL'19)*, pages 1–11. Springer, Cham.

Davis, L. W. and Metcalf, G. E. (2016). Does better information lead to better choices? evidence from energy-efficiency labels. *Journal of the Association of Environmental and Resource Economists*, 3(3):589–625.

Dubin, J. A. and McFadden, D. L. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica*, 52(2):345–362.

Eichler, H.-G., Abadie, E., Breckenridge, A., Flamion, B., Gustafsson, L. L., Leufkens, H., Rowland, M., Schneider, C. K., and Bloechl-Daum, B. (2011). Bridging the efficacy–effectiveness gap: a regulator's perspective on addressing variability of drug response. *Nature Reviews Drug Discovery*, 10(7):495–506.

Eizenberg, A. (2014). Upstream innovation and product variety in the U.S. home PC market. *The Review of Economic Studies*, 81(3):1003–1045.

European Commission (2019a). Commission regulation (eu) 2019/2023 of 1 october 2019 laying down ecodesign requirements for household washing machines and household washer-dryers pursuant to directive 2009/125/ec of the european parliament and of the council, amending commission regulation (ec) no 1275/2008 and repealing commission regulation (eu) no 1015/2010 (text with eea relevance). Official Journal of the European Union, OJ L 315, 5.12.2019, p. 285. EUR-Lex.

European Commission (2019b). Impact assessment accompanying the ecodesign and energy labelling regulations for household washing machines and household washer-dryers. Commission Staff Working Document SWD(2019) 349 final, European Commission.

European Commission (2021). European energy labels: rescaling and transition periods. https://commission.europa.eu/system/files/2021-04/rescaled_eu_energy_labels_and_transition_period.pdf. Guidance document on rescaled EU energy labels.

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the us daily newspaper market. *American Economic Review*, 103(5):1598–1628.

Fan, Y. and Yang, C. (2025). Estimating discrete games with many firms and many decisions: An application to merger and product variety. *Journal of Political Economy*, 133(6):1886–1931.

Goettler, R. L. and Gordon, B. R. (2011). Does amd spur intel to innovate more? *Journal of Political Economy*, 119(6):1141–1200.

Gupta, A. (2021). Impacts of performance pay for hospitals: The readmissions reduction program. *American Economic Review*, 111(4):1241–1283.

Hall, G. and Rust, J. (2003). Simulated minimum distance estimation of a model of optimal commodity price speculation with endogenously sampled prices. *manuscript, Yale University*.

Holmes, T. J. (2011). The diffusion of Wal-Mart and economies of density. *Econometrica*, 79(1):253–302.

Holmstrom, B. and Milgrom, P. (1991). Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design. *The Journal of Law, Economics, and Organization*, 7(special_issue):24–52.

Houde, J.-F., Newberry, P., and Seim, K. (2023). Nexus tax laws and economies of density in E-Commerce: A study of Amazon's fulfillment center network. *Econometrica*, 91(1):147–190.

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

Houde, S. (2026). The incidence of coarse certification: Evidence from the energy star program. *The RAND Journal of Economics*, 57(3):545–585.

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

Ito, K., Sallee, J. M., and Smith, J. A. (2026). Global policy spillovers: How environmental policies propagate through product attributes. Technical report, National Bureau of Economic Research.

Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the chicago public schools. *Journal of Public Economics*, 89(5–6):761–796.

Jia, P. (2008). What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica*, 76(6):1263–1316.

Josephy, B., Bush, E., Attali, S., and Zuloaga, F. R. (2015). Cold wash: The cool and modern way to launder. In *Proceedings of the 2015 ECEEE Summer Study on Energy Efficiency: First Fuel Now*, pages 1527–1535. European Council for an Energy Efficient Economy.

Knittel, C. R. (2011). Automobiles on steroids: Product attribute trade-offs and technological progress in the automobile sector. *American Economic Review*, 101(7):3368–3399.

Lu, S. F. (2012). Multitasking, information disclosure, and product quality: Evidence from nursing homes. *Journal of Economics & Management Strategy*, 21(3):673–705.

Montag, F. (2023). Mergers, foreign competition, and jobs: Evidence from the us appliance industry. *George J. Stigler Center for the Study of the Economy & the State Working Paper*, (326).

Reimers, I. and Waldfogel, J. (2021). Digitization and pre-purchase information: The causal and welfare impacts of reviews and crowd ratings. *American Economic Review*, 111(6):1944–1971.

Reynaert, M. (2021). Abatement strategies and the cost of environmental regulation: Emission standards on the european car market. *The Review of Economic Studies*, 88(1):454–488.

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

Rothwell, P. M. (2005). External validity of randomised controlled trials: "to whom do the results of this trial apply?". *The Lancet*, 365(9453):82–93.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *The RAND Journal of Economics*, 37(3):619–640.

Sinner, H. (1960). *Über das Waschen mit Haushaltswaschmaschinen*. Haus und Heim Verlag, Hamburg, Germany.

Su, R., Yang, C., and Sweeting, A. (2026). Competition, procurement and learning-by-doing in the space launch industry. Working Paper 34766, National Bureau of Economic Research.

Vatter, B. (2022). Quality disclosure and regulation: Scoring design in medicare advantage. *Available at SSRN 4250361*.

Wollmann, T. G. (2018). Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. *American Economic Review*, 108(6):1364–1406.

Appendix

Appendix Contents

A Regulatory Details **51**

A.1 Energy-Efficiency Index Calculations 51

A.2 Program-Design Constraints 53

B Data Imputation **57**

C Laundry-habits survey **59**

D Proofs and Extensions for Regulation-Induced Energy Shifting **64**

D.1 Proof of Proposition 1 64

D.2 Extension: Joint Eco Constraints 65

D.3 Socially Efficient Program Design 67

E Descriptive Evidence **68**

E.1 Energy Equation Estimation and Accounting 71

E.2 Matching Old-Label and Dual-Label Products 71

F Estimation and Computational Implementation **75**

F.1 Predicting Optimal Program Inputs and Profits 75

F.2 Reference Designs and Estimation Sample 77

F.3 Constrained Design Solver 78

G Conditional Portfolio Estimation: Theory, Validation, and Implementation **78**

G.1 Theoretical Results 78

G.2 Monte Carlo 83

G.3 Estimation Sample and Computational Implementation 87

G.4 Model Fit 89

H Counterfactual Setup and Computation **93**

H.1 Policy experiments 93

H.2 Market states and product outcomes 94

H.3 Entry equilibrium and outcome integration 94

I Welfare Accounting **95**

I.1 Purchase surplus and the label-energy correction 95

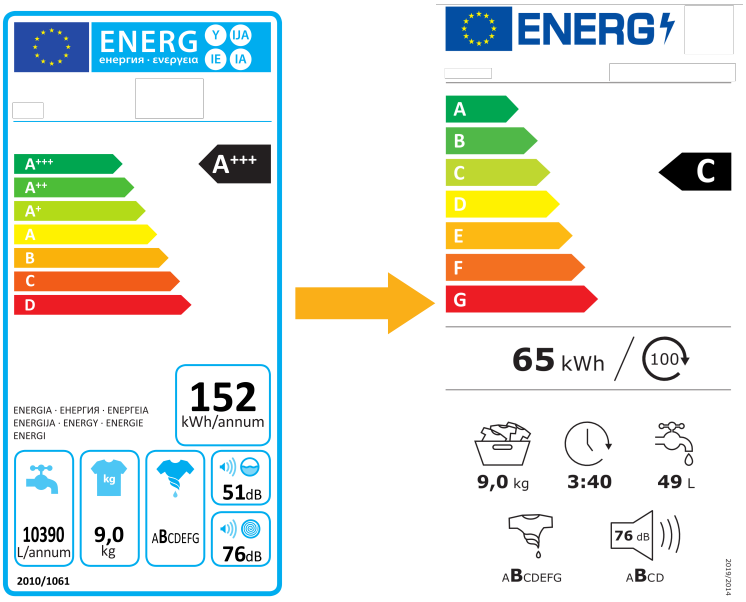
I.2 Electricity consumption and expenditure 97

I.3 Profits and reported comparisons 98

A Regulatory Details

This appendix summarizes the energy-efficiency calculations and program-design constraints under the 2011 and 2021 EU regulations for household washing machines.The old and new energy labels are shown below.

Figure A6: Energy Labels of Washing Machines in the EU



(a) 2011 Energy Label (b) 2021 Energy Label

Notes: The figure compares two versions of the EU energy label for washing machines. Both labels are shown for the same washing machine model and report the attributes of its energy-saving wash program. To preserve anonymity, brand and model identifiers are removed. The labels report energy-efficiency class and selected program attributes, including energy use, capacity, water consumption, program duration, and noise.

A.1 Energy-Efficiency Index Calculations

A.1.1 2011 Labeling Regime

Under the 2011 regime, the Energy Efficiency Index (EEI) measures weighted annual energy consumption relative to a capacity-based benchmark:

$$\mathrm{EEI}_{2011} = 100 \times \frac{\mathrm{AE}_c}{\mathrm{SAE}_c}, \qquad \mathrm{SAE}_c = 47c + 51.7, \tag{A.1}$$

where AE_c denotes annual energy consumption in kWh/year and c is the rated capacity in kg, defined as the lower of the full-load capacities of the standard Cotton 60°C and Cotton 40°C programs.

Annual energy consumption is calculated as

$$\mathrm{AE}_c = 220E_t + \frac{(P_o + P_l)(525600 - 200T_t)}{2 \times 60 \times 1000}, \tag{A.2}$$

where E_t is weighted energy consumption per cycle (kWh), P_o and P_l are weighted power consumption in off-mode and left-on mode (W), and T_t is weighted program duration (minutes). The calculation assumes 220 standard washing cycles per year and includes energy consumed in low-power modes.[26]

Weighted energy consumption is calculated from three standardized Cotton test cycles:

$$E_t = \frac{3E_{60,\mathrm{full}} + 2E_{60,\mathrm{half}} + 2E_{40,\mathrm{half}}}{7}, \tag{A.3}$$

where $E_{60,\mathrm{full}}$ and $E_{60,\mathrm{half}}$ denote energy consumption of the standard Cotton 60°C program at full and half load, respectively, and $E_{40,\mathrm{half}}$ denotes energy consumption of the standard Cotton 40°C program at half load. The same weights apply to P_o, P_l, and T_t.

The resulting EEI determines the energy-efficiency class according to the following thresholds.

Table A6: Energy efficiency classes

Energy efficiency class	**Energy Efficiency Index**
A+++ (most efficient)	$\mathrm{EEI} < 46$
A++	$46 \leq \mathrm{EEI} < 52$
A+	$52 \leq \mathrm{EEI} < 59$
A	$59 \leq \mathrm{EEI} < 68$

A.1.2 2021 Labeling Regime

Under the 2021 regime, the EEI is calculated from weighted energy consumption of the Eco 40–60 program:

$$\mathrm{EEI}_{2021} = 100 \times \frac{E_W}{\mathrm{SCE}_W}, \qquad \mathrm{SCE}_W = -0.0025c^2 + 0.0846c + 0.3920, \tag{A.4}$$

where E_W is weighted energy consumption in kWh per cycle and c is the rated capacity of the Eco 40–60 program in kg.

[26] The expression applies to the standard calculation without a separate adjustment for automatic power management. For machines equipped with a power-management system, the regulation provides an alternative formula incorporating the duration of left-on mode.

Weighted energy consumption combines measurements at full, half, and quarter load:

$$E_W = AE_{W,\text{full}} + BE_{W,\text{half}} + CE_{W,\text{quarter}}, \tag{A.5}$$

where the capacity-dependent weights for machines with rated capacity above 3 kg are

$$\begin{aligned} A &= -0.0391c + 0.6918, \\ B &= -0.0109c + 0.3582, \\ C &= 1 - A - B. \end{aligned} \tag{A.6}$$

For machines with rated capacity at or below 3 kg, the regulation assigns $A = 1$ and $B = C = 0$.

The resulting EEI determines the energy-efficiency class according to the following thresholds.

Table A7: Energy efficiency classes

Energy Efficiency Class	**Energy Efficiency Index**
A	$\text{EEI} \leq 52$
B	$52 < \text{EEI} \leq 60$
C	$60 < \text{EEI} \leq 69$
D	$69 < \text{EEI} \leq 80$
E	$80 < \text{EEI} \leq 91$
F	$91 < \text{EEI} \leq 102$
G	$\text{EEI} > 102$

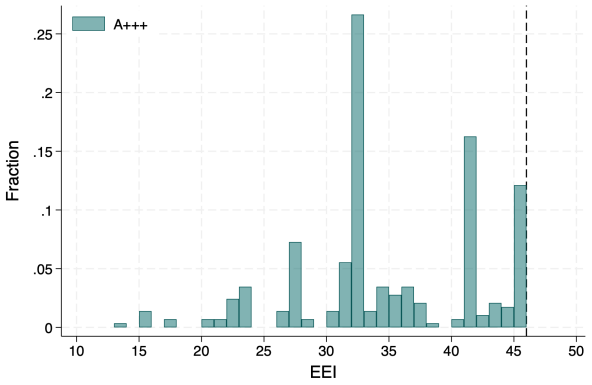
A.2 Program-Design Constraints

A.2.1 Cycle-Duration Limits under the 2021 Regime

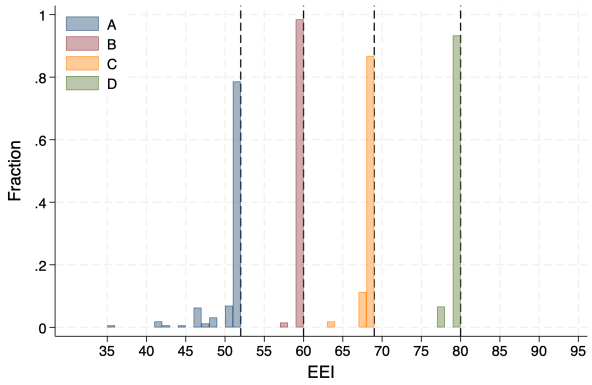
The 2021 ecodesign regulation imposes capacity-dependent upper bounds on the duration of the Eco 40–60 program. The limits differ between full and partial loads:

$$\begin{aligned} t_{\text{cap}}^{\text{full}}(c) &= \min\{137 + 10.2c,\ 240\}, \\ t_{\text{cap}}^{\text{half}}(c) = t_{\text{cap}}^{\text{quarter}}(c) &= \min\{120 + 6c,\ 180\}, \end{aligned} \tag{A.7}$$

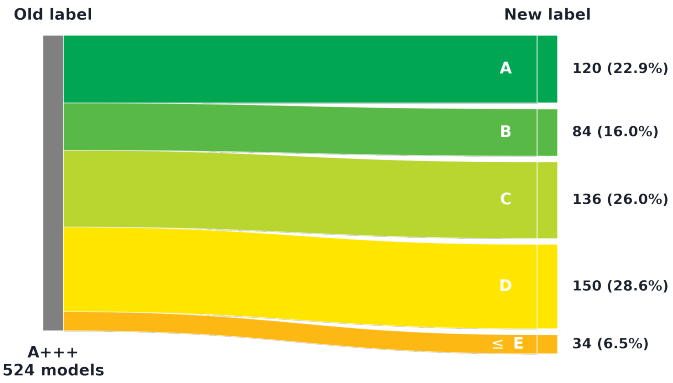
where c is rated capacity in kg and duration is measured in minutes. Table A8 reports the corresponding limits for common washing-machine capacities.



(a) Old label



(b) New label

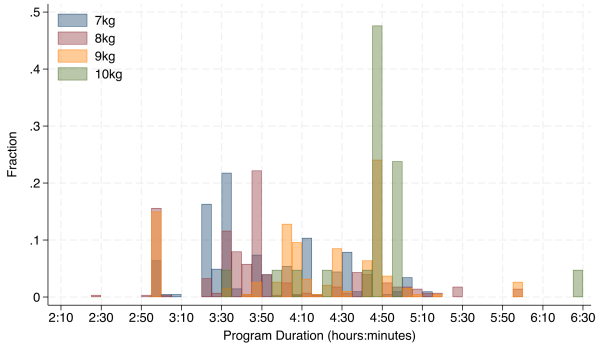


(c) Portfolio transformation

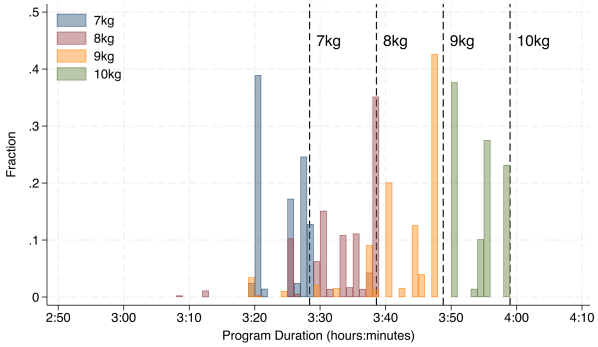
Figure A7: Energy Efficiency Indices and Portfolio Transformation

Notes: Panel (a) shows the distribution of EEI11 among A^{+++} products sold between March 2019 and December 2020. Panel (b) shows the distribution of EEI21 among products classified A–D under the new labeling regime. Panel (c) shows portfolio transformation across labeling regimes. Because the reform changed the underlying efficiency calculation and testing requirements, EEI11 and EEI21 are not directly comparable.

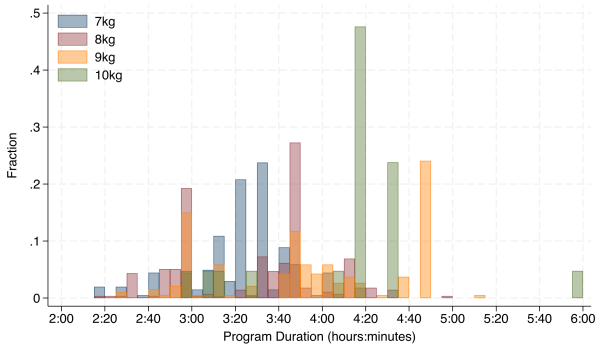
Figure A8: Eco program duration distributions under the old and new labels



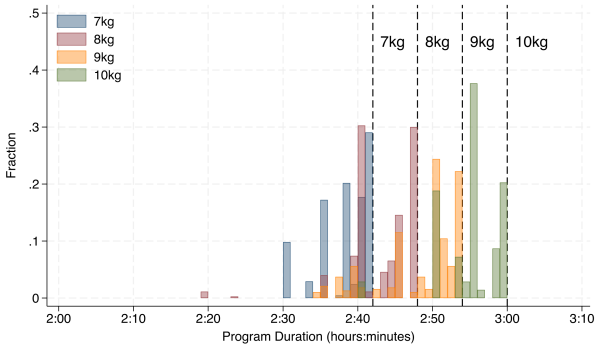
(a) Old label: Cotton 60°C, full load



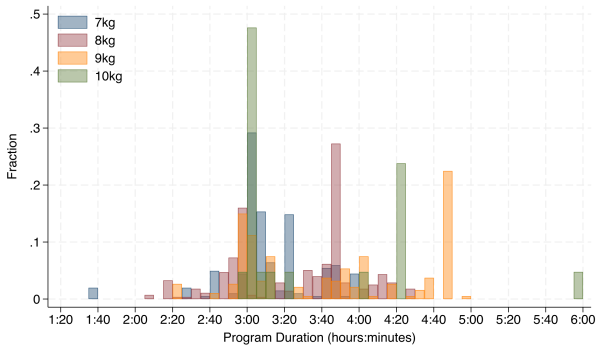
(b) New label: Eco, full load



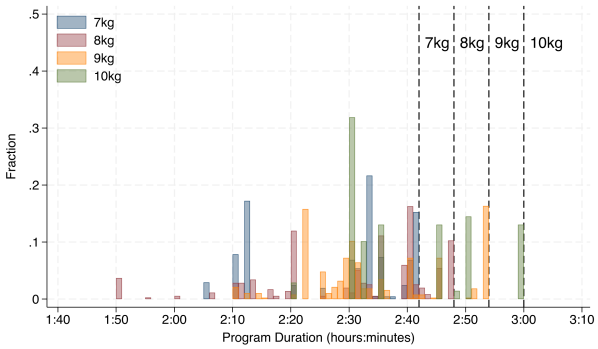
(c) Old label: Cotton 60°C, half load



(d) New label: Eco, half load



(e) Old label: Cotton 40°C, half load



(f) New label: Eco, quarter load

Notes: The figure shows program-duration distributions separately by rated capacity, with equal weight assigned to each model within each capacity group. The left column reports the three standard Cotton test conditions used under the old label; the right column reports Eco at full, half, and quarter load under the new label. Rows therefore do not always represent identical test conditions across regimes. Vertical dashed lines in the right column indicate capacity-specific regulatory duration limits. For plotting only, reported durations exceeding their applicable limits are adjusted to the largest whole minute strictly below the limit; the underlying data are unchanged.

Table A8: Eco 40–60 Duration Limits under the 2021 Regulation

Rated capacity	Full load	Half/quarter load
6 kg	198.2 min	156 min
7 kg	208.4 min	162 min
8 kg	218.6 min	168 min
9 kg	228.8 min	174 min
10 kg	239.0 min	180 min
≥10.5 kg	240 min	180 min

A.2.2 Water-Consumption Limits under the 2011 and 2021 Regimes

Under the 2011 ecodesign regulation, water-consumption limits applied to the standard Cotton 60°C program at full load. From December 1, 2011, the requirement was

$$W_t \leq 5c + 35, \tag{A.8}$$

where W_t is the water consumption of the standard Cotton 60°C program at full load (liters per cycle), and c is the lower of the rated capacities of the standard Cotton 60°C and Cotton 40°C programs at full load.

From December 1, 2013, the requirement was tightened to

$$W_t \leq 5c_{1/2} + 35, \tag{A.9}$$

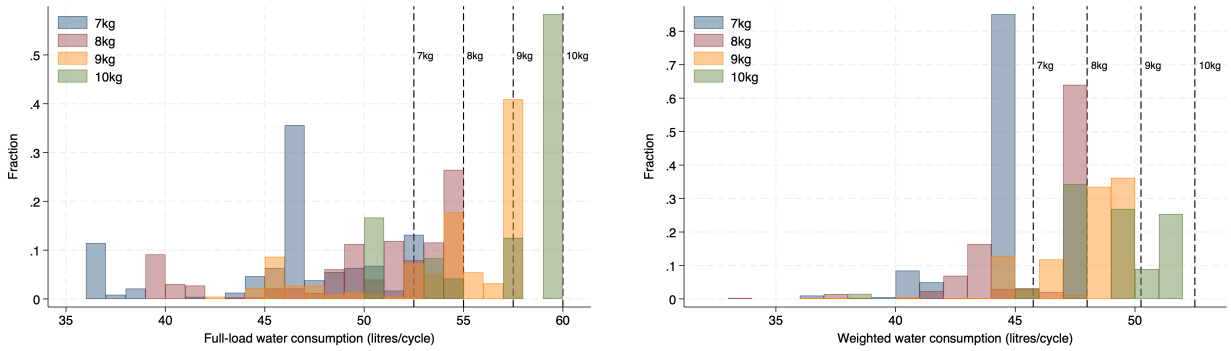
where $c_{1/2}$ is the lower of the rated capacities of the standard Cotton 60°C and Cotton 40°C programs at partial load.

The 2021 regulation instead constrains weighted water consumption of the Eco 40–60 program:

$$W_W \leq 2.25c + 30, \tag{A.10}$$

where W_W is weighted water consumption in liters per cycle, calculated across full, half, and quarter loads using the same capacity-dependent weights A, B, and C as in the energy-efficiency calculation.

Thus, the two regimes differ in both the regulated test programs and the measurement basis used to impose water-consumption limits.



(a) Old label: full-load water consumption

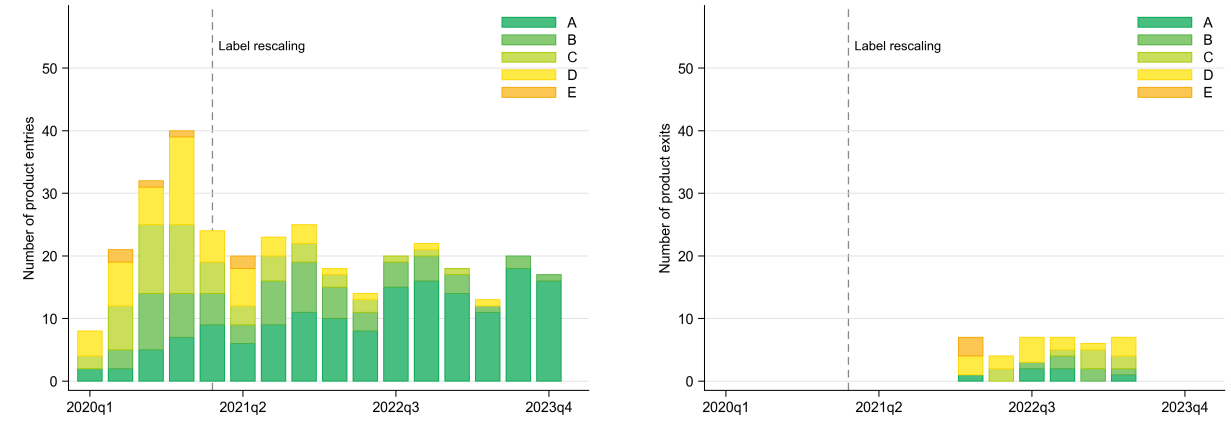
(b) New label: weighted water consumption

Figure A9: Distribution of Water Consumption

Notes: The figure shows water consumption in litres per cycle, separately by rated capacity, with equal weight assigned to each model within each capacity group. Panel (a) reports standard Cotton 60°C water consumption at full load under the old label. Panel (b) reports weighted Eco water consumption across full, half, and quarter loads under the new label. Vertical dashed lines indicate the capacity-specific regulatory limits. For plotting only, reported values exceeding their applicable limits are adjusted to the largest whole litre strictly below the limit. The two panels use different test conditions and aggregation rules and are therefore not directly comparable.

B Data Imputation

Figure B10: Product turnover by new energy class



(a) Entry

(b) Exit

Notes: Panel (a) plots quarterly product entries and panel (b) plots recorded exits for new-label classes *A*–*E*, 2020Q1–2023Q4. Products are counted once nationally. A product exits when all constituent models have recorded exits; the sample endpoint is not counted as an exit. Colors indicate new energy classes, and dashed lines mark the 2021Q1 label rescaling.

Standard-program temperature. Under the pre-2021 labeling regime, product manuals identify the regulated test cycles as the standard Cotton 60°C program at full and half load and the standard Cotton 40°C program at half load, but they generally do not report the temperature actually reached during these cycles. I do not assign the nominal program temperature to these observations. This distinction is empirically important because the nominal program name need not correspond to the realized wash temperature. Industry and consumer testing has documented substantial departures from the nominal temperature. In particular, Czyzewski (2022), writing from the perspective of BSH and the CENELEC working group on laundry appliances, reports that measured sump temperatures in Cotton 60°C programs vary substantially across products and that most machines do not reach 60°C; temperatures experienced by the laundry are lower still. An earlier Stiftung Warentest market test provides a quantitative benchmark: among 14 tested washing machines, none reached 60°C in the tested 60°C cotton program, with reported temperatures generally around 45–50°C and the lowest observation approximately 44°C (Van Holsteijn en Kemna et al., 2014).

I therefore impute realized temperatures from the physical characteristics of the wash cycle rather than from its nominal program name. I estimate an XGBoost regression using Eco 40–60 cycles for which temperature is observed. The estimation sample pools full-, half-, and quarter-load Eco cycles and excludes observations whose temperature was itself previously model-imputed. The predictors are program load, energy consumption, water consumption, program duration, product capacity, and the ratio of program load to rated capacity. All predictors are required to be observed and strictly positive. Importantly, the specification does not include the nominal program temperature, program identity, energy-label class, manufacturer, or the certified aggregate energy measure. The identifying assumption for the transfer is that, conditional on these physical inputs, the relationship between temperature and the other dimensions of program design is common across the Eco and standard test programs.

I evaluate the prediction rule using a product-level holdout sample: 20 percent of brand–model combinations are excluded from estimation, so that cycles from the same product cannot appear in both the training and validation samples. Within the Eco programs, the holdout regression yields an R^2 of 0.988, an RMSE of 0.79°C, and a mean absolute error of 0.42°C. I then re-estimate the model on all eligible Eco observations and apply it to the three old-regime standard programs using their observed physical inputs. The procedure imputes 19,521 standard-program temperatures; observations with incomplete or nonpositive physical inputs remain missing.

The resulting temperatures are well below the nominal program settings. The median imputed temperature is 36.1°C for the standard Cotton 60°C full-load cycle, 34.7°C for the standard Cotton 60°C half-load cycle, and 32.1°C for the standard Cotton 40°C half-load cycle. These estimates are lower than the roughly 45–50°C range reported in the 2013 Stiftung Warentest test, but both

the external tests and the imputation imply the same economically important feature of the old testing regime: the nominal 40°C or 60°C designation is a poor measure of the temperature actually delivered by an energy-efficient cycle. The lower imputed values should therefore not be interpreted as reproducing the Stiftung Warentest distribution mechanically; rather, the external evidence supports the central feature of the imputation—substantial temperature reductions relative to the nominal setting.

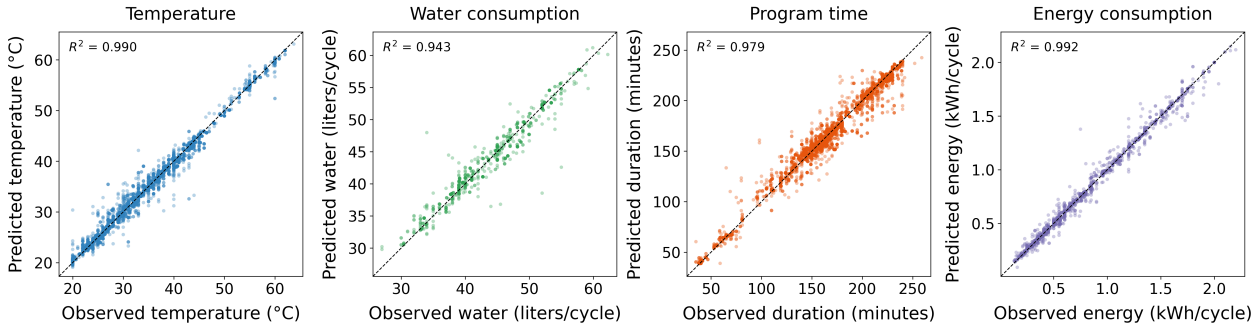


Figure B11: Program inputs out-of-sample fit

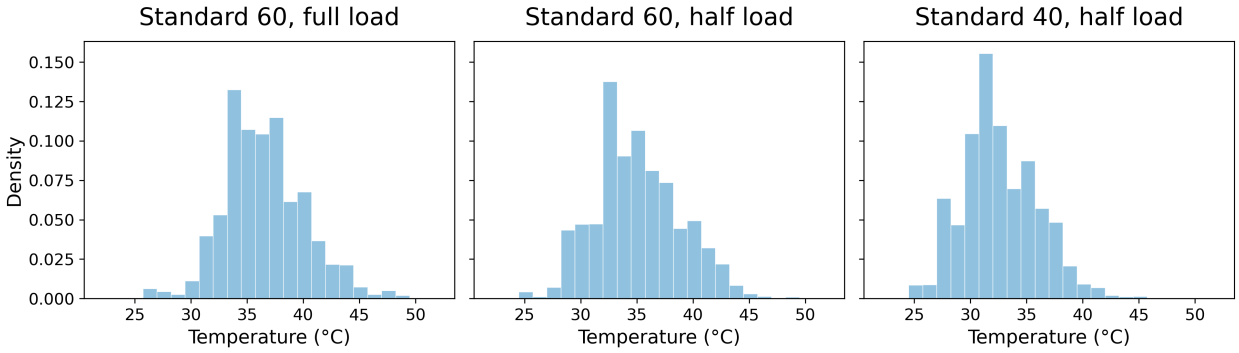


Figure B12: Distribution of standard program temperature

C Laundry-habits survey

Dear Sir or Madam,

This questionnaire has been designed by the University of Wisconsin–Madison (USA). It is intended for: (1) consumers who own a washing machine with the post-2021 A–G energy label, and (2) the primary person responsible for doing the laundry in the household. Completing the questionnaire will take approximately 5–10 minutes.

1. Who is mainly responsible for doing the laundry in your household? Please choose one option.

- ☐ I am mainly responsible.
- ☐ I share responsibility with someone else.
- ☐ Someone else in the household is mainly responsible.

2. In which federal state (Bundesland) do you live? Please choose one option.

   - ☐ Baden-Württemberg
   - ☐ Bavaria (Bayern)
   - ☐ Berlin
   - ☐ Brandenburg
   - ☐ Bremen
   - ☐ Hamburg
   - ☐ Hesse (Hessen)
   - ☐ Lower Saxony (Niedersachsen)
   - ☐ Mecklenburg-Vorpommern
   - ☐ North Rhine-Westphalia (Nordrhein-Westfalen)
   - ☐ Rhineland-Palatinate (Rheinland-Pfalz)
   - ☐ Saarland
   - ☐ Saxony (Sachsen)
   - ☐ Saxony-Anhalt (Sachsen-Anhalt)
   - ☐ Schleswig-Holstein
   - ☐ Thuringia (Thüringen)

3. How many people are living in your household, including yourself? Please choose one option.

   - ☐ 1 person
   - ☐ 2 people
   - ☐ 3 people
   - ☐ 4 people
   - ☐ More than 4 people

4. Approximately what is your household's monthly income?

5. What is the **brand and model number** of your washing machine?
   (*You can usually find the model number on the energy label or on the* Typenschild */ rating plate—often inside the door opening/door frame, on the back of the machine, or in the user manual/receipt. For Bosch/Siemens, it may be listed as the "E-Nr."; for AEG/Electrolux, it may be the "PNC".*)

   _____

6. What is the energy efficiency class of your washing machine (*A*–*G* label)? Please choose one option.

   ☐ A
   ☐ B
   ☐ C
   ☐ D
   ☐ E
   ☐ I'm not sure.

7. What is the rated load capacity of your washing machine (in kg)? Please choose one option.

   ☐ 6 kg or less
   ☐ 7 kg
   ☐ 8 kg
   ☐ 9 kg
   ☐ 10 kg or more
   ☐ I'm not sure.

8. When did you purchase your washing machine? Please choose one option.

2021:	☐Q1	☐Q2	☐Q3	☐Q4
2022:	☐Q1	☐Q2	☐Q3	☐Q4
2023:	☐Q1	☐Q2	☐Q3	☐Q4
2024:	☐Q1	☐Q2	☐Q3	☐Q4
2025:	☐Q1	☐Q2	☐Q3	☐Q4

9. When purchasing a washing machine, which information do you use to judge energy efficiency? Please choose the option that best describes you.

☐ I use the energy efficiency class (*A*–*G*) as a quick filter.

☐ I mainly compare the energy consumption values (kWh per 100 cycles) on the energy label and choose the lower value.

☐ I look at both the class (*A*–*G*) and the kWh value, but I do not calculate electricity costs.

☐ I convert the kWh value into an electricity cost estimate.

☐ I rarely or never use the energy efficiency class or the kWh value.

☐ I'm not sure.

10. *Only if you selected in the previous question that you convert the kWh value into an electricity cost estimate:*
    Approximately how much do you think your washing machine costs in electricity per year? Please choose one option.

    ☐ Less than €10

    ☐ €10–€20

    ☐ €21–€30

    ☐ €31–€40

    ☐ €41–€60

    ☐ €61–€80

    ☐ More than €80

    ☐ I'm not sure.

11. What would you expect a washing machine with the highest energy efficiency class to save on? Please choose one option.

    ☐ Energy in all programmes

    ☐ Energy and water in all programmes

    ☐ Energy in some selected programmes

    ☐ Energy and water in some selected programmes

12. How often do your laundry loads have the following stain levels? *"Stains" include visible marks (e.g., food, grease, mud) and/or strong odours (e.g., sweat). Please choose one option per row.*

	Not at all	**Once per 4 weeks**	**Once per 3 weeks**	**Once per 2 weeks**	**1–2x per week**	**3–4x per week**	**5–6x per week**	**> 6x per week**
Low	☐	☐	☐	☐	☐	☐	☐	☐
Medium	☐	☐	☐	☐	☐	☐	☐	☐
High	☐	☐	☐	☐	☐	☐	☐	☐

Low stains: No visible stains; lightly worn; no strong odour.

Medium stains: Some visible dirt/marks or noticeable odour (typical everyday laundry).

High stains: Multiple items with heavy/visible stains (e.g., mud/grease/food) and/or strong odour.

13. How often do you use the following washing programmes for washing your clothes? Please choose one option per row.

	Not at all	Once per 4 weeks	Once per 3 weeks	Once per 2 weeks	1–2x per week	3–4x per week	5–6x per week	>6x per week
Cotton 20 °C / Cold wash	☐	☐	☐	☐	☐	☐	☐	☐
Cotton 30 °C	☐	☐	☐	☐	☐	☐	☐	☐
Cotton 40 °C	☐	☐	☐	☐	☐	☐	☐	☐
Cotton 60 °C	☐	☐	☐	☐	☐	☐	☐	☐
Cotton 90 °C	☐	☐	☐	☐	☐	☐	☐	☐
Synthetic / Easy care 30 °C / 40 °C	☐	☐	☐	☐	☐	☐	☐	☐
Eco 40-60 (Eco) / Energy saving	☐	☐	☐	☐	☐	☐	☐	☐
Silk/Wool	☐	☐	☐	☐	☐	☐	☐	☐
Quick wash / Short	☐	☐	☐	☐	☐	☐	☐	☐
Mix (all fabrics)	☐	☐	☐	☐	☐	☐	☐	☐
Other	☐	☐	☐	☐	☐	☐	☐	☐

D Proofs and Extensions for Regulation-Induced Energy Shifting

D.1 Proof of Proposition 1

Under logit program choice, expected utility is

$$\mathcal{I}(d_E, d_C) = \log\left(e^{\beta d_E} + e^{\beta d_C}\right).$$

Let s_r denote the corresponding program-choice probability. Differentiating yields

$$\frac{\partial \mathcal{I}}{\partial d_E} = \beta s_E, \qquad \frac{\partial \mathcal{I}}{\partial d_C} = \beta s_C, \qquad \frac{\partial^2 \mathcal{I}}{\partial d_C \partial d_E} = -\beta^2 s_E s_C.$$

Product demand is $Q(z)$, where

$$z = P + \kappa e_E(d_E) - \theta \mathcal{I}(d_E, d_C),$$

with $Q' < 0$, $\kappa > 0$, and $\theta > 0$. Throughout, partial derivatives of Q with respect to program inputs refer to the composite demand function $Q(z)$. Applying the chain rule gives

$$\frac{\partial^2 Q}{\partial d_C \partial d_E} = \theta\beta s_C \left[\beta s_E \left(\theta Q''(z) + Q'(z)\right) - \kappa e_E'(d_E) Q''(z)\right].$$

The sign depends on the curvature of product demand and the direct contribution of Eco energy consumption. Proposition 1 imposes $\frac{\partial^2 Q}{\partial d_C \partial d_E} < 0$ as a sufficient condition for decreasing differences in demand.

With the Eco-duration cap binding, $d_E^* = \bar{d}_E$. The Cotton first-order condition is

$$\frac{\partial Q}{\partial d_C} = C_C'(d_C).$$

Implicit differentiation with respect to $\bar{d}_E$ yields

$$\frac{\partial d_C^*}{\partial \bar{d}_E} = \frac{\frac{\partial^2 Q}{\partial d_C \partial d_E}}{C_C'' - \frac{\partial^2 Q}{\partial d_C^2}} < 0,$$

where the inequality follows from decreasing differences in demand and the strict second-order condition $C_C'' - \frac{\partial^2 Q}{\partial d_C^2} > 0$.

Because Cotton energy consumption satisfies $e_C'(d_C) > 0$,

$$\frac{\partial e_C(d_C^*)}{\partial \bar{d}_E} = e_C'(d_C^*) \frac{\partial d_C^*}{\partial \bar{d}_E} < 0.$$

Thus, tightening the Eco-duration cap increases optimal Cotton duration and energy consumption, establishing Proposition 1. □

D.2 Extension: Joint Eco Constraints

I extend the stylized model to allow firms to choose a second Eco input x_E, such as temperature, while Cotton duration remains its only design input. Eco energy consumption is $e_E(d_E, x_E)$, with $\frac{\partial e_E}{\partial d_E} > 0$ and $\frac{\partial e_E}{\partial x_E} > 0$. Consumer utility from Eco becomes $v_E = \beta d_E + \beta_x x_E + \varepsilon_E$, while Cotton utility remains $v_C = \beta d_C + \varepsilon_C$. Expected program utility is

$$\mathcal{I}(d_E, x_E, d_C) = \log\left(e^{\beta d_E + \beta_x x_E} + e^{\beta d_C}\right).$$

Taking price as given and normalizing the markup to one, the firm solves

$$\begin{aligned} \max_{d_E, x_E, d_C} \quad & Q\left(P + \kappa e_E(d_E, x_E) - \theta \mathcal{I}(d_E, x_E, d_C)\right) - C_E(d_E, x_E) - C_C(d_C) \\ \text{s.t.} \quad & d_E \le \bar{d}_E, \qquad e_E(d_E, x_E) \le \bar{e}. \end{aligned}$$

Proposition 2 (Energy Shifting under Joint Eco Constraints)**.** *Suppose both Eco constraints bind and Cotton duration has an interior local optimum satisfying the second-order condition. If* $\frac{\partial^2 Q}{\partial d_C \partial d_E} < 0$ *and* $\frac{\partial^2 Q}{\partial d_C \partial x_E} < 0$*, tightening the Eco-energy cap increases optimal Cotton duration and energy consumption, holding the duration cap fixed. Jointly tightening both caps also increases Cotton duration and energy consumption if the second Eco input does not increase.*

Proof. Under logit program choice, the cross-partials of expected utility are

$$\frac{\partial^2 \mathcal{I}}{\partial d_C \partial d_E} = -\beta^2 s_E s_C, \qquad \frac{\partial^2 \mathcal{I}}{\partial d_C \partial x_E} = -\beta \beta_x s_E s_C.$$

As in Proposition 1, the corresponding demand cross-partials also depend on the curvature of Q and the direct contribution of Eco energy consumption. The proposition assumes both are negative.

Because both constraints bind,

$$d_E^* = \bar{d}_E, \qquad e_E(\bar{d}_E, x_E^*) = \bar{e}.$$

Differentiating the energy constraint gives

$$\mathrm{d}x_E^* = \frac{\mathrm{d}\bar{e} - \frac{\partial e_E}{\partial d_E}\mathrm{d}\bar{d}_E}{\frac{\partial e_E}{\partial x_E}}. \tag{D.1}$$

The Cotton first-order condition is

$$\frac{\partial Q}{\partial d_C} = C_C'(d_C).$$

Implicit differentiation yields

$$\left(C_C'' - \frac{\partial^2 Q}{\partial d_C^2}\right) \mathrm{d}d_C^* = \frac{\partial^2 Q}{\partial d_C \partial d_E} \mathrm{d}\bar{d}_E + \frac{\partial^2 Q}{\partial d_C \partial x_E} \mathrm{d}x_E^*.$$

Substituting equation (D.1) gives

$$\mathrm{d}d_C^* = \frac{1}{C_C'' - \frac{\partial^2 Q}{\partial d_C^2}} \left[\left(\frac{\partial^2 Q}{\partial d_C \partial d_E} - \frac{\frac{\partial e_E}{\partial d_E}}{\frac{\partial e_E}{\partial x_E}} \frac{\partial^2 Q}{\partial d_C \partial x_E} \right) \mathrm{d}\bar{d}_E + \frac{\frac{\partial^2 Q}{\partial d_C \partial x_E}}{\frac{\partial e_E}{\partial x_E}} \mathrm{d}\bar{e} \right]. \tag{D.2}$$

Holding the duration cap fixed,

$$\frac{\partial d_C^*}{\partial \bar{e}} = \frac{\frac{\partial^2 Q}{\partial d_C \partial x_E}}{\frac{\partial e_E}{\partial x_E}\left(C_C'' - \frac{\partial^2 Q}{\partial d_C^2}\right)} < 0.$$

The inequality follows from decreasing differences between Cotton duration and the second Eco input, increasing Eco energy consumption in x_E, and the strict second-order condition. Thus, tightening the energy cap reduces x_E and increases optimal Cotton duration.

Now consider jointly tightening both caps, with $\mathrm{d}\bar{d}_E < 0$ and $\mathrm{d}\bar{e} < 0$. Equation (D.1) implies that the second Eco input does not increase if

$$\mathrm{d}\bar{e} \leq \frac{\partial e_E}{\partial d_E} \mathrm{d}\bar{d}_E.$$

Under this condition, $\mathrm{d}d_E^* < 0$ and $\mathrm{d}x_E^* \leq 0$. Because both demand cross-partials are negative, the Cotton first-order condition implies $\mathrm{d}d_C^* > 0$.

Finally, since $e_C'(d_C) > 0$, any increase in Cotton duration also increases Cotton energy consumption. This establishes both claims. □

Holding the energy cap fixed instead gives

$$\left.\frac{\partial d_C^*}{\partial \bar{d}_E}\right|_{\bar{e}} = \frac{\frac{\partial^2 Q}{\partial d_C \partial d_E} - \frac{\frac{\partial e_E}{\partial d_E}}{\frac{\partial e_E}{\partial x_E}} \frac{\partial^2 Q}{\partial d_C \partial x_E}}{C_C'' - \frac{\partial^2 Q}{\partial d_C^2}}.$$

The sign is ambiguous because tightening the duration cap permits an offsetting increase in the second Eco input. Joint constraints strengthen energy shifting when they prevent such compensation within Eco.

D.3 Socially Efficient Program Design

I extend the stylized model to compare private and socially efficient program designs. Consumers perceive Eco energy consumption as representative of both programs. To isolate the resulting information distortion, suppose the social cost of electricity equals its valuation at purchase. Since θ converts program utility into price-equivalent units, define $\eta = \kappa/\theta$.

Under consumers' beliefs, expected program utility net of perceived electricity costs is

$$\mathcal{I}^P(d_E, d_C) = \mathcal{I}(d_E, d_C) - \eta e_E(d_E).$$

A social planner instead accounts for actual program-specific electricity consumption:

$$\mathcal{I}^S(d_E, d_C) = \log\left(e^{\beta d_E - \eta e_E(d_E)} + e^{\beta d_C - \eta e_C(d_C)}\right).$$

The socially efficient Cotton usage share is

$$s_C^S = \frac{e^{\beta d_C - \eta e_C(d_C)}}{e^{\beta d_E - \eta e_E(d_E)} + e^{\beta d_C - \eta e_C(d_C)}}.$$

Thus, private and socially efficient program choices differ by

$$\log\frac{s_C}{s_E} - \log\frac{s_C^S}{s_E^S} = \eta\left[e_C(d_C) - e_E(d_E)\right].$$

When Cotton consumes more electricity than Eco, private choices allocate too many cycles to Cotton relative to the social benchmark.

To characterize socially efficient design, let

$$z^P = P + \kappa e_E(d_E) - \theta\mathcal{I}(d_E, d_C), \qquad z^S = P - \theta\mathcal{I}^S(d_E, d_C).$$

Let consumer surplus under full information be

$$CS(z^S) = \int_{z^S}^{\infty} Q(u)\, du.$$

Maintaining the unit-markup normalization and abstracting from other distortions, the social planner chooses program durations to maximize total surplus:

$$\max_{d_E, d_C} \quad CS(z^S) + Q(z^S) - C_E(d_E) - C_C(d_C) \qquad \text{s.t.} \qquad d_E \le \bar{d}_E.$$

This objective includes consumer surplus, operating profits, and design costs, with actual electricity costs incorporated into $\mathcal{I}^S$.

The socially optimal Cotton duration satisfies

$$\theta \left[Q(z^S) - Q'(z^S) \right] s_C^S \left[\beta - \eta \frac{\partial e_C}{\partial d_C} \right] = C_C'(d_C^S). \tag{D.3}$$

By contrast, the firm's private optimum satisfies

$$-\theta Q'(z^P) \beta s_C = C_C'(d_C^P). \tag{D.4}$$

The social planner accounts for the additional electricity consumed when Cotton duration increases. The firm does not face this direct energy penalty because consumers perceive Cotton energy consumption as equal to Eco energy consumption. However, the two objectives also differ in their valuation of product demand and consumer surplus.

To compare the optimal designs, hold Eco duration fixed at the binding regulatory cap. Suppose the social objective is strictly concave in Cotton duration. Evaluated at the private optimum d_C^P, a sufficient condition for excessive Cotton duration is

$$\left[Q(z^S) - Q'(z^S) \right] s_C^S \left[\beta - \eta \frac{\partial e_C}{\partial d_C} \right] < -Q'(z^P) \beta s_C. \tag{D.5}$$

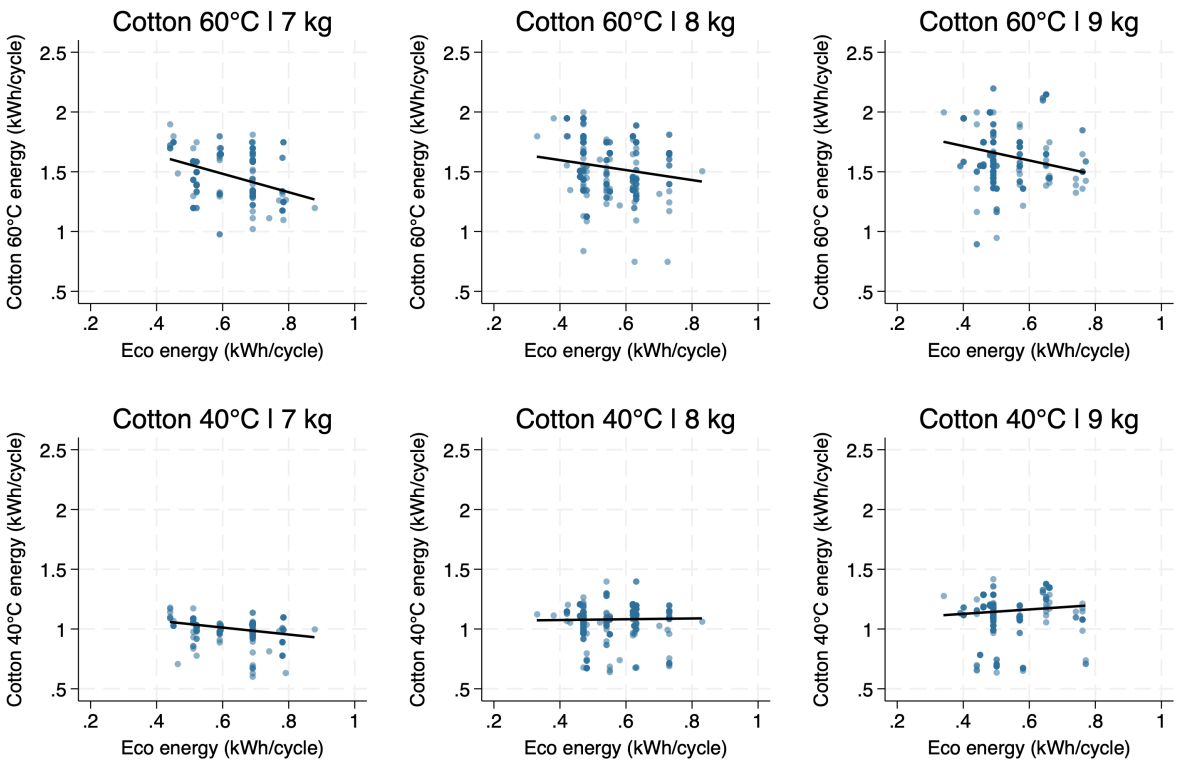
Under this condition, the social marginal return to extending Cotton is smaller than the private marginal return at d_C^P. Strict concavity then implies

$$d_C^S < d_C^P, \qquad e_C(d_C^S) < e_C(d_C^P).$$

For example, when $\beta > 0$, the condition is satisfied if $\eta \frac{\partial e_C}{\partial d_C} \geq \beta$ at the private optimum: the social marginal electricity cost of extending Cotton is at least as large as its marginal duration benefit. More generally, whether private designs are excessively energy intensive depends on the relative marginal benefits in equation (D.5).

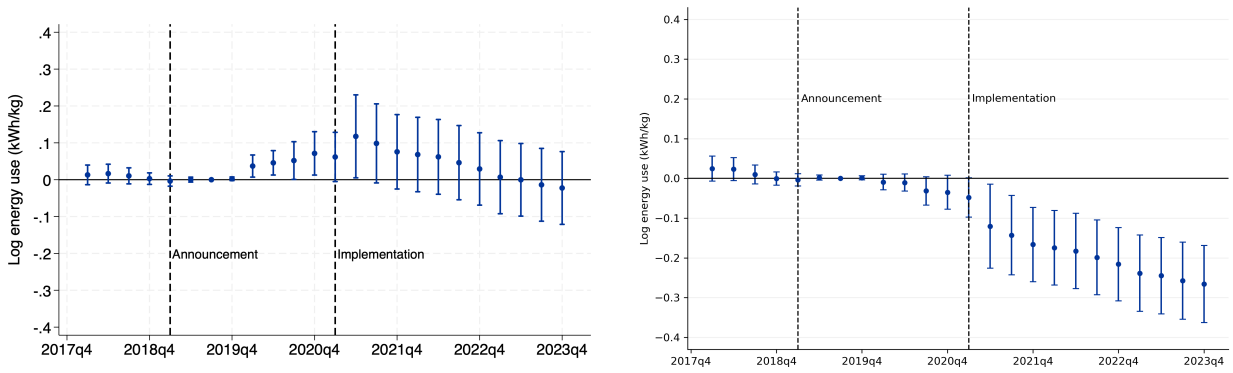
E Descriptive Evidence

Figure E1: Eco and Cotton Energy Consumption by Capacity



Notes: Each point represents a washing-machine model carrying the new energy label. The horizontal axis reports certified Eco energy consumption, and the vertical axis reports Cotton energy consumption, both in kWh per cycle. The top and bottom rows show Cotton 60°C and Cotton 40°C, respectively; columns distinguish capacities of 7, 8, and 9 kg. Solid lines show separate unadjusted linear fits within each panel. The sample includes models with nonmissing energy consumption for both programs, including imputed Cotton values.

Figure E2: Changes in Eco Energy Consumption: Alternative Measures



(a) Eco: alternative definition

(b) Eco: label-reported consumption

Notes: Both panels report changes in log Eco energy consumption per kilogram relative to Wool using alternative consumption measures. Panel (a) combines Standard 60°C at full and half load for the old program, retaining the normalized weights and new-program definition from the main specification. Panel (b) uses label-reported consumption under the respective old and new testing protocols. Dots denote coefficients from equation (1), normalized to zero in 2019Q3. The sample contains 31 product lines (firm × capacity × spin speed), observed in every quarter from 2018Q1 to 2023Q4, with products weighted equally within each line. Bars denote 95% confidence intervals based on standard errors clustered by product line. Dashed vertical lines mark the reform's announcement in March 2019 and implementation in March 2021.

E.1 Energy Equation Estimation and Accounting

I estimate equation (2) by unweighted OLS with brand–model fixed effects and program-group effects for Eco, Cotton, and Easy Care. Old and new Eco form the omitted group; Cotton 20°C, Cotton 40°C, and Cotton 60°C share one group effect. The sample includes Eco observations with imputed values, while Cotton and Easy Care observations require nonimputed energy consumption and all four inputs. Wool is excluded. Temperature is measured in degrees Celsius, load in kilograms, duration in hours, and water use in liters. Identification requires the disturbance to have conditional mean zero given the inputs and fixed effects. Standard errors are conventional homoskedastic OLS standard errors, accounting for absorbed product effects. The robustness specification retains only old and new Eco observations, with product fixed effects and no separate old/new Eco indicator.

For the accounting exercise, products receive equal weights within each quarter. Eco outcomes use common normalized full- and half-load weights across labeling regimes; quarter load is excluded. Product turnover and imputed inputs are included. Table A2 reports percentage changes in arithmetic means. The contribution calculations instead use changes in mean logarithms relative to the reference quarter, denoted by 0:

$$C_{T,t} = \widehat{\theta}_T \left(\overline{\log T}_t - \overline{\log T}_0 \right), \qquad C_{L,t} = (\widehat{\theta}_L - 1) \left(\overline{\log L}_t - \overline{\log L}_0 \right),$$
$$C_{d,t} = \widehat{\theta}_d \left(\overline{\log d}_t - \overline{\log d}_0 \right), \qquad C_{w,t} = \widehat{\theta}_w \left(\overline{\log w}_t - \overline{\log w}_0 \right).$$

The load coefficient is adjusted because the outcome is energy consumption per kilogram. The conditional-efficiency component is $C_{z,t} = \bar{z}_t - \bar{z}_0$. Within each program group, the group fixed effect cancels over time.

Reported contribution shares equal

$$S_{k,t} = 100 \frac{C_{k,t}}{\sum_{h \in \{T,L,d,w,z\}} |C_{h,t}|}.$$

Positive shares indicate contributions toward higher energy consumption per kilogram; negative shares indicate contributions toward lower consumption. The denominator includes the unreported conditional-efficiency component, so the absolute shares displayed in the table need not sum to 100 percent. These shares describe the relative magnitudes of fitted log contributions, not fractions of the observed change in arithmetic mean energy consumption. Residual and aggregation effects are excluded

E.2 Matching Old-Label and Dual-Label Products

For each dual-label product, I construct a comparison group of old-label-only products with the same brand and exactly matching old Eco full-load duration and laundry load. Old-label-only

Table A1: Energy Equation Estimates

	Main specification		Old and new Eco only	
	Estimate	Std. error	Estimate	Std. error
Log load	0.167	0.008	0.175	0.008
Log temperature	1.580	0.007	1.703	0.015
Log duration	0.401	0.013	0.258	0.014
Log water	0.321	0.015	0.306	0.015
Product fixed effects	Yes		Yes	
Program-group fixed effects	Yes		No	
Observations	8,990		6,390	
Products	1,786		1,786	
Adjusted R^2	0.936		0.956	

Notes: The dependent variable is log electricity consumption per cycle (kWh). Both specifications use unweighted OLS with model fixed effects. The main specification includes Eco, Cotton, and Easy Care group effects, with old and new Eco pooled as the reference group. Cotton 20°C, Cotton 40°C, and Cotton 60°C share one group effect. The robustness specification retains only old and new Eco observations, without a separate old/new Eco indicator. Standard errors are conventional homoskedastic OLS standard errors, accounting for absorbed product effects.

products have an observed old energy class but no new energy class. Matching uses the old Eco 60°C full-load program and includes imputed attributes where available.

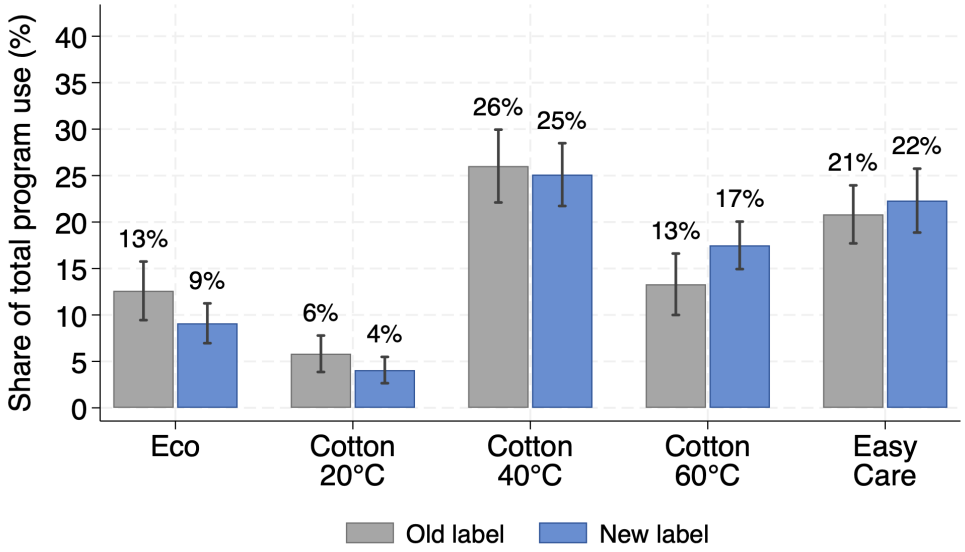
For each target program—Cotton 40°C, Cotton 60°C, or Easy Care—I calculate the arithmetic mean of each outcome among matched old-label-only products. This mean defines $\widehat{q}_{jr}^{\text{old}}$. The dependent variable is the dual-label product's log outcome minus the log of this matched mean. Each regression retains observations with positive, nonmissing outcomes and a valid matched mean, so sample sizes may differ across programs and outcomes. These comparisons measure differences from matched products rather than within-product changes over time.

Table A2: Changes in Program Inputs and Their Contributions to Energy Consumption

	Energy/kg	Load		Duration		Temperature		Water	
Quarter	Δ%	Δ%	Contrib. (%)	Δ%	Contrib. (%)	Δ%	Contrib. (%)	Δ%	Contrib. (%)
Panel A: Eco									
2020Q2	1.6	0.0	-0.8	-1.0	-14.7	0.9	71.3	0.8	11.1
2020Q4	5.8	0.6	-5.1	-3.2	-12.2	3.6	67.6	3.3	12.1
2021Q2	16.1	1.7	-5.0	-9.1	-11.6	10.4	64.1	10.1	12.2
2021Q4	10.0	2.6	-9.0	-9.0	-12.6	7.9	54.1	9.3	12.5
2022Q2	5.8	3.7	-13.7	-8.4	-12.5	5.9	42.8	10.7	14.6
2022Q4	1.7	2.9	-13.8	-8.5	-16.7	3.1	28.4	10.4	17.9
2023Q2	-2.8	3.8	-23.5	-8.1	-20.4	0.3	-0.5	11.6	25.3
2023Q4	-4.5	3.9	-20.4	-8.0	-17.1	-0.7	-12.2	12.2	22.5
Panel B: Cotton 60°C									
2020Q2	0.9	0.0	-2.2	1.1	34.9	-0.2	-30.3	0.9	28.3
2020Q4	3.4	0.6	-11.8	4.9	37.5	-0.4	-15.2	3.8	29.4
2021Q2	12.4	1.9	-10.8	14.0	33.0	-0.8	-11.1	12.9	31.3
2021Q4	14.3	2.9	-14.7	16.0	31.4	-1.0	-11.0	12.4	25.7
2022Q2	14.1	4.1	-18.9	17.5	30.7	-0.4	-4.5	13.8	25.6
2022Q4	16.9	3.2	-15.3	18.3	33.7	0.1	0.1	14.1	27.6
2023Q2	16.2	4.3	-18.2	18.5	30.5	0.6	4.6	14.7	25.6
2023Q4	15.9	4.3	-17.9	18.2	29.2	0.8	5.7	15.0	25.3
Panel C: Cotton 40°C									
2020Q2	0.9	0.2	-6.9	1.4	34.5	0.3	23.5	0.9	17.5
2020Q4	4.0	0.6	-8.1	5.7	36.7	1.0	25.0	3.9	20.5
2021Q2	11.1	1.3	-6.0	16.4	39.6	3.0	29.6	11.9	23.2
2021Q4	8.8	2.4	-11.0	16.4	36.8	3.2	29.2	11.0	19.9
2022Q2	9.0	3.4	-13.3	18.1	34.2	3.5	26.7	12.6	19.3
2022Q4	9.8	3.0	-11.3	18.1	33.4	3.6	27.3	12.8	18.9
2023Q2	9.5	4.0	-14.2	18.2	31.9	3.6	25.8	13.4	18.8
2023Q4	9.2	4.3	-15.3	17.9	31.0	3.5	24.8	13.5	18.7
Panel D: Easy Care									
2020Q2	1.0	-0.1	11.6	1.1	37.5	-0.1	-12.3	0.4	11.7
2020Q4	4.6	-0.1	1.7	4.8	36.1	1.2	40.3	1.6	10.3
2021Q2	13.5	-0.1	1.1	14.0	34.2	4.6	52.8	5.0	10.2
2021Q4	12.2	-0.2	2.6	12.8	34.0	3.9	49.7	4.4	9.9
2022Q2	12.9	1.0	-3.7	14.6	32.3	4.5	47.0	4.7	8.9
2022Q4	14.0	0.8	-2.5	15.0	30.4	5.0	47.9	5.1	8.7
2023Q2	13.2	2.4	-8.5	15.9	28.4	5.4	45.5	5.0	7.5
2023Q4	12.7	2.9	-10.5	15.6	27.2	5.3	43.9	5.1	7.5

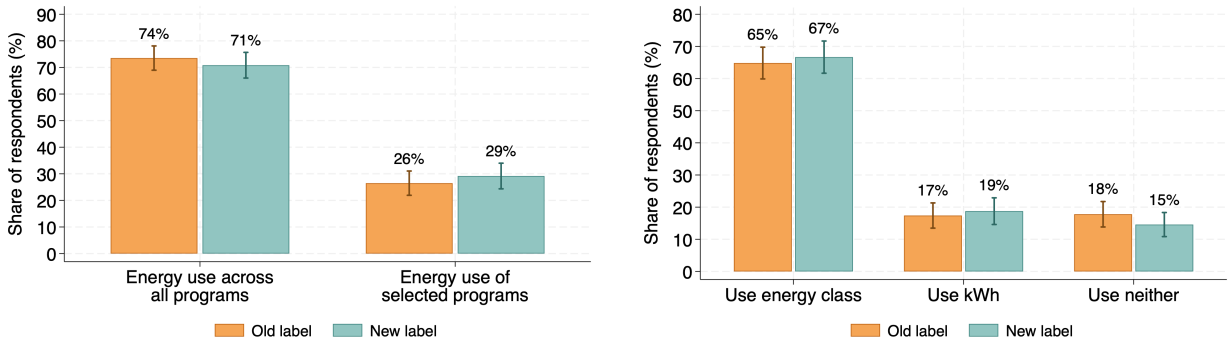
Notes: Changes in arithmetic means are relative to 2019Q4. Contribution shares equal $100C_k/\sum_h |C_h|$, using changes in mean log inputs and estimated coefficients; load uses $\widehat{\theta}_L - 1$. The denominator includes the unreported conditional-efficiency component, so displayed absolute shares need not sum to 100. Residual and aggregation effects are excluded. Products are equally weighted. Eco uses common normalized full- and half-load weights, excluding quarter load. Product turnover and imputed inputs are included.

Figure E3: Share of program use



Notes: The figure reports the distribution of wash-program choices in the household survey, separately for owners of old- and new-label washing machines. Shares are calculated as the fraction of total reported wash cycles assigned to each program. Eco accounts for 13 percent of cycles among old-label households and 9 percent among new-label households. Cotton 40°C, Cotton 60°C, and Easy Care are the most frequently used programs. Remaining cycles belong to other programs not shown in the figure.

Figure E4: Consumer Beliefs About Energy Ratings and Energy Costs



(a) Interpretation of Energy Ratings

(b) Information Used to Assess Energy Costs

Notes: The figure reports household survey responses separately for owners of old- and new-label washing machines. Panel (a) shows whether respondents believe the energy rating reflects energy consumption across all programs or only selected programs. Panel (b) reports whether respondents use the energy class, reported kWh consumption, or neither to assess energy costs. Most respondents in both groups interpret the energy rating as measuring whole-machine efficiency, while approximately two-thirds rely on the energy class rather than reported kWh consumption.

F Estimation and Computational Implementation

F.1 Predicting Optimal Program Inputs and Profits

Entry estimation requires evaluating firms' continuation payoffs under many alternative product portfolios. To avoid repeatedly solving the program-design game, I train a neural network to predict equilibrium program inputs and firm profits net of design costs. The network follows the DeepSets architecture of Zaheer et al. (2017), allowing it to accommodate portfolios of different sizes without making predictions depend on the ordering of products.

Training Data. The training data combine the observed portfolio path with 1,500 simulated paths covering 2020Q1–2023Q4. Each simulated path specifies entry decisions across product lines over the 16-quarter horizon. Total introductions range from 150 to 450 across paths. Conditional on this total, I sample line-quarter entry slots without replacement, allowing at most one introduction per line-quarter. Entrants receive classes A, B, C, and D with probabilities $(0.1, 0.2, 0.3, 0.4)$ during 2020 and $(0.4, 0.3, 0.2, 0.1)$ thereafter. These draws generate variation in portfolio size, energy-class composition, and competition.

Each path begins with the pre-existing product portfolio. Legacy products follow their prescribed exits, while simulated entrants remain available through the end of the horizon. I solve each path chronologically, carrying previously established designs into subsequent quarters. Firms choose program inputs when first introducing an energy class on a line; repeated introductions inherit that class's existing design. The resulting equilibrium inputs, operating profits, and profits net of design costs provide the training targets. Only solutions that pass the numerical feasibility and convergence checks enter the training sample.

Product Features and Portfolio Aggregation. Let x_{jt} collect the information available before program design for product j in quarter t. Product features include firm ownership, capacity, spin speed, energy class, labeling regime, predetermined conditional efficiency, and fixed program attributes. They also include regional demand fixed effects and market sizes, calendar and transition-period indicators, and the information governing design choices: reference inputs, adjustment-cost coefficients, input bounds, engineering offsets, regulatory parameters, and indicators for adjustable inputs. The inputs to be chosen are masked, so their optimized values enter only as prediction targets.

A shared encoder maps each product's features into a vector,

$$h_{jt} = \phi_\vartheta(x_{jt}). \tag{F.1}$$

The network then summarizes these vectors within the firm's portfolio, across rival products, and across the market. Let $\mathcal{J}_{ft}$ denote firm f's portfolio, $n_{ft} = |\mathcal{J}_{ft}|$, and $n_t = |\mathcal{J}_t|$. The three summaries are

$$\begin{aligned} H_{ft}^{\text{own}} &= \frac{1}{n_{ft}} \sum_{j \in \mathcal{J}_{ft}} h_{jt}, \\ H_{ft}^{\text{rival}} &= \frac{1}{n_t - n_{ft}} \sum_{j \in \mathcal{J}_t \setminus \mathcal{J}_{ft}} h_{jt}, \\ H_t^{\text{market}} &= \frac{1}{n_t} \sum_{j \in \mathcal{J}_t} h_{jt}. \end{aligned} \tag{F.2}$$

An empty group receives a zero vector. Because averaging does not preserve the number of products, I also include portfolio-size measures,

$$b_{ft} = \big(\log(1 + n_{ft}),\, \log(1 + n_t)\big). \tag{F.3}$$

Together, these summaries describe the firm's own offerings and its competitive environment without depending on how products are ordered.

Prediction and Training. For product j owned by firm f, the network combines its product vector with the own-firm and rival summaries to predict program inputs:

$$\widehat{q}_{jt} = \underline{q}_{jt} + (\overline{q}_{jt} - \underline{q}_{jt}) \odot \sigma\left[g_\vartheta^q\left(h_{jt}, H_{ft}^{\text{own}}, H_{ft}^{\text{rival}}, b_{ft}\right)\right], \tag{F.4}$$

where σ is the elementwise logistic function and $\odot$ denotes elementwise multiplication. This transformation keeps predicted inputs within their bounds. Predetermined and inherited inputs have identical lower and upper bounds and therefore remain fixed.

A separate output function predicts operating profits and profits net of design costs:

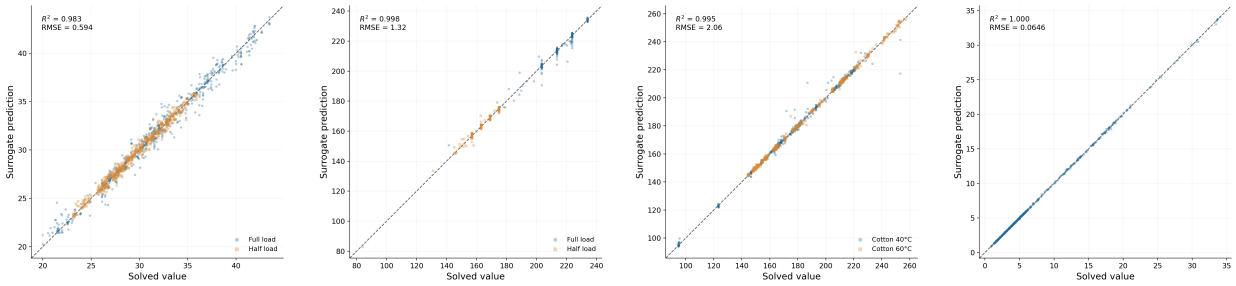
$$\begin{pmatrix} \widehat{\pi}_{ft} \\ \widehat{W}_{ft} \end{pmatrix} = g_\vartheta^W\left(H_{ft}^{\text{own}}, H_t^{\text{market}}, H_{ft}^{\text{rival}}, b_{ft}\right). \tag{F.5}$$

Reordering products leaves firm-profit predictions unchanged and simply reorders the corresponding product-input predictions.

The shared encoder has two hidden layers of width 96 with hyperbolic-tangent activations, and each output function has one hidden layer of the same width. Training minimizes scaled squared errors for adjustable inputs and the two profit targets. Features are standardized using the training sample. Approximately 70 percent of path groups are allocated to training, 15 percent to validation, and 15 percent to testing. All quarters and policy variants of a path remain in the

same group, preventing information from a shared portfolio history from appearing across splits. Figure F5 reports out-of-sample fit on the held-out test paths, comparing predicted program inputs and operating profits with their numerical solutions.

Figure F5: Surrogate Fit: Optimal Program Inputs and Profits



(a) Eco temperature (b) Eco duration (c) Cotton duration (d) Operating profit

Notes: Horizontal axes show numerical solutions; vertical axes show surrogate predictions. Panels (a) and (b) distinguish full- and half-load Eco programs, while panel (c) distinguishes Cotton 40°C and Cotton 60°C. Temperatures are measured in degrees Celsius, durations in minutes, and profits in millions of euros. Dashed lines indicate equality. Each panel reports R^2 and root mean squared error (RMSE).

F.2 Reference Designs and Estimation Sample

Reference designs determine the point from which firms incur input-adjustment costs. For duration, I use the matched pre-reform A^{+++} design on the same product line, including fitted baseline Cotton durations where available. This reference captures the line's established cycle structure: changing total duration may require reorganizing the timing of washing, rinsing, and other control sequences. Keeping this benchmark fixed treats subsequent duration choices as departures from that established design. For Eco temperature, I instead use the most recent earlier first introduction of an energy class on the line, averaging across simultaneous introductions when necessary. This reference allows temperature adjustments to build on the line's previous design as firms move across efficiency classes. These are modeling assumptions about adjustment costs rather than separately identified engineering restrictions. Conditional energy efficiency, water use, and other predetermined inputs remain fixed.

The estimation sample includes first introductions of an energy class on a product line, for which the model permits redesign. Repeated introductions inherit the class's existing design and do not provide separate input-adjustment targets. Fixed coordinates are excluded from the corresponding moments, and full- and half-load Eco observations are pooled within each input group. Estimation uses fixed reference data constructed from observed design histories. In simulated paths, references are updated using the path's own previously optimized designs.

F.3 Constrained Design Solver

For each candidate design-cost vector, I solve the program-design game conditional on the product portfolio and predetermined attributes. Given rivals' designs q_{-f}, firm f chooses its adjustable inputs to maximize expected operating profits net of design costs:

$$q_f^{\mathrm{BR}}(q_{-f}) \in \arg\max_{q_f \in \mathcal{Q}_f} \left\{ \mathbb{E}_{\xi,\omega}\big[\pi_f(q_f, q_{-f}; \xi, \omega)\big] - \sum_{j \in \mathcal{J}_f^{\mathrm{red}}} DC_j(q_j) \right\}, \tag{F.6}$$

where $\mathcal{Q}_f$ incorporates the applicable EEI and duration constraints, and operating profits are evaluated at the subsequent pricing equilibrium. Only inputs associated with first introductions are optimized; incumbent and inherited designs remain fixed.

I solve each firm's constrained problem using sequential least squares programming (SLSQP) (Kraft, 1988). At each step, SLSQP forms a quadratic approximation to the objective and linear approximations to the constraints, then updates the candidate input vector. To solve the design game, I update firms sequentially, using the latest available rival designs. Iteration continues until designs stabilize and no firm can obtain a material improvement through the numerical best-response search. The resulting profile q^* approximately satisfies

$$q_f^* \in q_f^{\mathrm{BR}}(q_{-f}^*) \qquad \text{for every firm } f. \tag{F.7}$$

I then use the resulting inputs to construct the model-implied estimation moments. Because the constrained problems need not be globally concave, these checks establish numerical convergence and local optimality rather than global optimality or uniqueness.

G Conditional Portfolio Estimation: Theory, Validation, and Implementation

This appendix establishes the theoretical foundations of conditional portfolio estimation, evaluates its performance in Monte Carlo experiments, and describes its implementation in the washing-machine model.

G.1 Theoretical Results

The results establish equilibrium equivalence between joint and compatible conditional probabilities, equality of the population identified sets under full and conditional likelihood, and first-order equivalence between exact and simulated conditional estimation. They apply to the incomplete-information portfolio game and impose conditions on response probabilities rather than a particular shock distribution.

Setup. Conditional on the public state X, firm f chooses $a_f \in \mathcal{A}_f = \prod_{\ell=1}^{L_f} \mathcal{A}_{f\ell}$, where each coordinate set is finite and nonempty. Private shocks ε_f are independent across firms conditional on X; dependence within a firm's shock vector is unrestricted. Let β collect the model parameters, $\mathcal{P} = (\mathcal{P}_f)_{f=1}^F$ denote the profile of portfolio distributions, and $\mathcal{P}_{-f}$ denote the induced product distribution of rivals' portfolios. Define the response map by

$$\Psi_{f,\beta}(\mathcal{P}, X)(a_f) = \Pr_{\varepsilon_f | X;\beta} \left\{ a_f = \arg\max_{b_f \in \mathcal{A}_f} \mathbb{E}_{\mathcal{P}_{-f}} \left[u_f(b_f, a_{-f}, \varepsilon_f, X; \beta) \right] \right\}, \tag{G.1}$$

with a fixed measurable tie-breaking rule. An equilibrium satisfies $\mathcal{P} = \Psi_\beta(\mathcal{P}, X)$. Write $\mathcal{Q} = \prod_f \Delta(\mathcal{A}_f)$ for the product of closed probability simplices and $\mathcal{Q}^+$ for profiles assigning positive probability to every portfolio.

For $\mathcal{P} \in \mathcal{Q}^+$, let $C(\mathcal{P})$ collect all within-firm full conditionals:

$$C(\mathcal{P})_{f\ell}(d \mid b) = \frac{\mathcal{P}_f(d, b)}{\sum_{e \in \mathcal{A}_{f\ell}} \mathcal{P}_f(e, b)}, \qquad b \in \mathcal{A}_{f,-\ell}. \tag{G.2}$$

Define $\mathcal{C} = C(\mathcal{Q}^+)$. Thus, $\mathcal{C}$ contains only compatible conditional arrays. The reconstruction result of Brook (1964) implies that each array in $\mathcal{C}$ determines a unique positive joint distribution for every firm. Denote the inverse of C by R.

Equilibrium equivalence. Fix (β, X) and define the conditional response map

$$T_\beta(c) = C\big(\Psi_\beta(R(c), X)\big), \qquad c \in \mathcal{C}. \tag{G.3}$$

Theorem 1 (Equilibrium equivalence)**.** *Suppose $\Psi_\beta(\cdot, X) : \mathcal{Q} \to \mathcal{Q}^+$ is continuous. Then T_β has a fixed point, and C and R give a one-to-one correspondence between the joint and conditional equilibrium sets:*

$$c = T_\beta(c) \quad \Longleftrightarrow \quad R(c) = \Psi_\beta(R(c), X). \tag{G.4}$$

On the fixed finite product support, for $\mathcal{P}^n, \mathcal{P} \in \mathcal{Q}^+$,

$$\max_f d_{\mathrm{TV}}(\mathcal{P}_f^n, \mathcal{P}_f) \longrightarrow 0 \quad \Longleftrightarrow \quad \|C(\mathcal{P}^n) - C(\mathcal{P})\|_\infty \longrightarrow 0. \tag{G.5}$$

For any fixed $c \in \mathcal{C}$, a Gibbs sampler that updates every coordinate of firm f in each systematic sweep converges to $R_f(c)$.

Proof. The set $\mathcal{Q}$ is nonempty, compact, and convex. Brouwer's fixed-point theorem gives $\mathcal{P}^\star = \Psi_\beta(\mathcal{P}^\star, X)$, and positivity places $\mathcal{P}^\star$ in $\mathcal{Q}^+$.

Since $\Psi_\beta(R(c), X) \in \mathcal{Q}^+$, the map T_β takes $\mathcal{C}$ into itself. Applying C to a joint fixed point gives a conditional fixed point. Conversely, applying R to $c = T_\beta(c)$ gives $R(c) = \Psi_\beta(R(c), X)$. The identities $R \circ C = \mathrm{id}$ and $C \circ R = \mathrm{id}$ establish the bijection and conditional existence. No convexity of $\mathcal{C}$ is required.

For equation (G.5), C is continuous at every positive joint distribution by equation (G.2). To verify continuity of its inverse, fix a reference portfolio a_f^0. Brook's reconstruction identity gives

$$\begin{aligned} r_f(a_f; c) &= \prod_{\ell=1}^{L_f} \frac{c_{f\ell}(a_{f\ell} \mid a_{f,<\ell}, a^0_{f,>\ell})}{c_{f\ell}(a^0_{f\ell} \mid a_{f,<\ell}, a^0_{f,>\ell})}, \\ R_f(c)(a_f) &= \frac{r_f(a_f; c)}{\sum_{b_f \in \mathcal{A}_f} r_f(b_f; c)}. \end{aligned} \tag{G.6}$$

All denominators are positive at $C(\mathcal{P})$, so the finite products and normalization are continuous there. On a finite support, componentwise convergence of joint probabilities is equivalent to total-variation convergence, proving equation (G.5).

Finally, each Gibbs coordinate update preserves $R_f(c)$, and strict positivity makes the transition matrix of a complete sweep positive. Hence $R_f(c)$ is the unique invariant distribution, and the chain converges to it from any initial portfolio (Casella and George, 1992). □

The theorem concerns compatible probabilities and a positive limiting distribution. It does not establish convergence of the outer equilibrium iteration or a mixing bound uniform in portfolio dimension. Gibbs sampling avoids explicit reconstruction when the full conditionals can be evaluated directly. Strict positivity and continuity must hold under the maintained shock specification.

Identification equivalence. Let $\mathcal{P}^\star(\beta, X)$ be a measurable selected equilibrium family and $c^\star(\beta, X) = C(\mathcal{P}^\star(\beta, X))$. Both criteria maintain this same family. For an observed portfolio profile Y, define

$$\begin{aligned} Q_F(\beta) &= \mathbb{E}\left[\sum_f \log \mathcal{P}^\star_f(Y_f \mid X; \beta)\right], \\ Q_C(\beta) &= \mathbb{E}\left[\sum_{f,\ell} w_{f\ell} \log c^\star_{f\ell}(Y_{f\ell} \mid Y_{f,-\ell}, X; \beta)\right], \qquad w_{f\ell} > 0. \end{aligned} \tag{G.7}$$

The second criterion is a composite conditional likelihood (Besag, 1975; Varin et al., 2011). Define

$$\mathcal{B}_0 = \left\{\beta : \mathcal{P}^\star_f(\cdot \mid X; \beta) = \mathcal{P}^\star_f(\cdot \mid X; \beta_0) \text{ for every } f, \text{ almost surely in } X\right\}. \tag{G.8}$$

Theorem 2 (Identification equivalence)**.** *Suppose the conditional distribution of the data is generated by* $\prod_f \mathcal{P}_f^\star(\cdot \mid X; \beta_0)$*. All candidate firm distributions are strictly positive on the same finite product support, the weights in equation* (G.7) *are fixed and positive, and both criteria are finite at every candidate parameter. Then*

$$\arg\max_\beta Q_F(\beta) = \arg\max_\beta Q_C(\beta) = \mathcal{B}_0. \tag{G.9}$$

Proof. Write $\mathcal{P}_f^0 = \mathcal{P}_f^\star(\cdot \mid X; \beta_0)$ and $c_{f\ell}^0$ for its full conditionals. The Kullback–Leibler decompositions give

$$Q_F(\beta_0) - Q_F(\beta) = \sum_f \mathbb{E}_X D_{\mathrm{KL}}\Big(\mathcal{P}_f^0 \,\|\, \mathcal{P}_f^\star(\cdot \mid X; \beta)\Big), \tag{G.10}$$

$$Q_C(\beta_0) - Q_C(\beta) = \sum_{f,\ell} w_{f\ell} \mathbb{E}_{X, Y_{f,-\ell}} D_{\mathrm{KL}}\Big(c_{f\ell}^0(\cdot \mid Y_{f,-\ell}, X) \,\|\, c_{f\ell}^\star(\cdot \mid Y_{f,-\ell}, X; \beta)\Big). \tag{G.11}$$

Both expressions are nonnegative. Equation (G.10) is zero exactly on $\mathcal{B}_0$.

If equation (G.11) is zero, every conditional divergence vanishes almost surely because every weight is positive. At almost every X, positivity of the true joint distribution assigns positive probability to every conditioning portfolio. The candidate and true full conditionals therefore agree for all firms, coordinates, and conditioning portfolios. Brook's reconstruction then gives equality of the joint distributions, so $\beta \in \mathcal{B}_0$. Conversely, equality of the joint distributions implies equality of their conditionals. Thus, both criteria have the same maximizer set. □

This is a population identification result within a correctly specified equilibrium family. It does not imply equal finite-sample estimates or asymptotic variances, and it does not determine equilibrium selection.

Simulation-valid inference. Consider M independent, identically distributed games, with fixed parameter dimension and finite action sets as $M \to \infty$. Let $U_m(\beta)$ be the conditional estimator's game-level estimating function, evaluated using exact equilibrium beliefs and expectations, and write

$$\overline{U}_M(\beta) = M^{-1} \sum_{m=1}^{M} U_m(\beta).$$

Any estimated first-stage parameters and their estimating equations are included in β and U_m. For an equilibrium-profiled likelihood, U_m is the total derivative of the game-level conditional criterion. Let $\widetilde{U}_M$ denote the corresponding numerical equations, including errors from Gibbs sampling, rival integration, and equilibrium computation.

Theorem 3 (Simulation-valid inference)**.** *Let $\widehat{\beta}_C$ and $\widehat{\beta}_G$ converge in probability to the same interior parameter β_0 on the same differentiable equilibrium branch, with*

$$\|\overline{U}_M(\widehat{\beta}_C)\| = o_p(M^{-1/2}), \qquad \|\widetilde{U}_M(\widehat{\beta}_G)\| = o_p(M^{-1/2}).$$

Suppose $\mathbb{E}[U_m(\beta_0)] = 0$, $\sqrt{M}\,\overline{U}_M(\beta_0) \Rightarrow N(0, \Omega)$, and, for every consistent sequence β_M,

$$\overline{U}_M(\beta_M) = \overline{U}_M(\beta_0) + J(\beta_M - \beta_0) + o_p\Big(M^{-1/2} + \|\beta_M - \beta_0\|\Big), \tag{G.12}$$

where J is nonsingular. If, on a fixed neighborhood $\mathcal{N}$ of β_0,

$$\sup_{\beta \in \mathcal{N}} \|\widetilde{U}_M(\beta) - \overline{U}_M(\beta)\| = o_p(M^{-1/2}), \tag{G.13}$$

then

$$\sqrt{M}(\widehat{\beta}_G - \widehat{\beta}_C) = o_p(1), \qquad \sqrt{M}(\widehat{\beta}_G - \beta_0) \Rightarrow N\Big(0, J^{-1}\Omega J^{-\top}\Big). \tag{G.14}$$

Proof. Consistency places $\widehat{\beta}_G$ in $\mathcal{N}$ with probability approaching one. Its numerical residual and equation (G.13) imply

$$\overline{U}_M(\widehat{\beta}_G) = \widetilde{U}_M(\widehat{\beta}_G) + \big[\overline{U}_M(\widehat{\beta}_G) - \widetilde{U}_M(\widehat{\beta}_G)\big] = o_p(M^{-1/2}).$$

Thus, both estimators are approximate roots of the same exact equations. For $j \in \{C, G\}$, set $d_j = \widehat{\beta}_j - \beta_0$. If $s_J > 0$ is the smallest singular value of J, the central limit condition and equation (G.12) give

$$s_J\|d_j\| \le O_p(M^{-1/2}) + o_p\Big(M^{-1/2} + \|d_j\|\Big).$$

Absorbing the term proportional to $\|d_j\|$ establishes $d_j = O_p(M^{-1/2})$. Substituting this rate into equation (G.12) yields

$$\sqrt{M}(\widehat{\beta}_j - \beta_0) = -J^{-1}\sqrt{M}\,\overline{U}_M(\beta_0) + o_p(1), \qquad j \in \{C, G\}.$$

Subtracting proves first-order equivalence, and the central limit theorem gives the stated distribution. This follows the standard estimating-equation expansion in van der Vaart (1998, Chapter 5). □

Under the usual differentiability and moment conditions, $J = \mathbb{E}[\partial U_m(\beta_0)/\partial\beta^\top]$ and $\Omega = \mathrm{Var}(U_m(\beta_0))$. The derivative includes equilibrium feedback. In independent simplex coordinates, if the response map is continuously differentiable and $I - \Psi_{\mathcal{P}}$ is nonsingular at the selected equilibrium, the implicit function theorem gives

$$\frac{\partial \mathcal{P}^\star}{\partial \beta^\top} = (I - \Psi_{\mathcal{P}})^{-1}\Psi_\beta. \tag{G.15}$$

Here, $\Psi_{\mathcal{P}}$ and Ψ_β denote the Jacobians with respect to beliefs and parameters. Consistent estimates $\widehat{J}$ and $\widehat{\Omega}$ yield $M^{-1}\widehat{J}^{-1}\widehat{\Omega}\widehat{J}^{-\top}$ for the estimator covariance, with scores aggregated at the game level.

Equation (G.13) restricts the combined numerical error in the estimating equations. Convergence of a fixed-target Gibbs chain alone does not establish this condition, and consistent covariance estimation requires separate control of derivative and second-moment approximations. The theorem compares simulated and exact conditional estimation; it does not equate their variance with that of full maximum likelihood.

G.2 Monte Carlo

The Monte Carlo experiments evaluate the accuracy and computational cost of conditional entry estimation with Gibbs sampling relative to full-portfolio estimation. Both estimators use a common nested pseudo-likelihood procedure.

Choice sets and sample size. Each replication consists of $M = 500$ independent markets, each with three firms. Each firm has L product lines and chooses $a_{f\ell} \in \{0, 1, 2\}$ on each line, corresponding to no entry or entry with one of two product types. A firm's portfolio belongs to $\mathcal{A}_L = \{0, 1, 2\}^L$. I consider $L \in \{3, 4, 5\}$ and conduct 100 replications for each specification. Table A3 reports the resulting sizes of the own-portfolio, rival-portfolio, and market action spaces.

Table A3: Size of the action space

Lines per firm	Own portfolios	Rival configurations	Market configurations
L	3^L	3^{2L}	3^{3L}
3	27	729	19,683
4	81	6,561	531,441
5	243	59,049	14,348,907

Notes: Each market contains three firms. Each firm has L product lines and chooses among three actions on each line: no entry or one of two products. A firm's portfolio therefore belongs to an action space of size 3^L. For each own portfolio, direct evaluation of expected profits sums over 3^{2L} joint portfolios of the two rival firms. The complete market action space contains 3^{3L} configurations.

Entry costs and operating profits. For each firm–line pair in market m, I independently draw an observed entry-cost shifter $z_{mf\ell} \sim U[-1, 1]$. The cost of portfolio a_f is

$$C_{mf}(a_f; \theta) = \sum_{\ell=1}^{L} \left[\mathbf{1}\{a_{f\ell} > 0\}(\kappa + \beta_z z_{mf\ell}) + \mathbf{1}\{a_{f\ell} = 2\}\delta_2 \right], \tag{G.16}$$

where $\theta_0 = (\kappa, \beta_z, \delta_2)' = (1.10, 0.45, 0.35)'$. Choosing no entry on a line incurs zero cost.

Each potential product has a positive scalar attribute. For each firm f, line ℓ, and entry action $a \in \{1, 2\}$, I independently draw $e_{f\ell a} \sim N(0, 1)$ once and set

$$\begin{aligned} &x_{f\ell}(0) = 0, \\ &x_{f\ell}(a) = \exp\{0.10(f-1) + 0.15(\ell-1) + 0.35(a-1) + 0.20 e_{f\ell a}\}, \qquad a \in \{1, 2\}. \end{aligned} \tag{G.17}$$

These attributes are public and fixed across markets and replications. Product type 2 has a higher expected attribute than type 1, although this ordering need not hold on every line. Define the firm's portfolio attribute as

$$X_f(a_f) = \sum_{\ell=1}^{L} x_{f\ell}(a_{f\ell}). \tag{G.18}$$

Operating profits are

$$\pi_f(a_f, a_{-f}) = D \frac{X_f(a_f)}{1 + X_f(a_f) + \sum_{g \neq f} X_g(a_g)}, \qquad D = 20. \tag{G.19}$$

An empty portfolio earns zero operating profit. Entry costs are subtracted once from expected operating profits.

Firms observe all public primitives and privately observe independent standard type-I extreme-value shocks for their complete portfolios, including the empty portfolio. These shocks are independent across portfolios, firms, and markets. Operating-profit primitives are known to the economist; only entry-cost parameters are estimated.

Holding rivals' portfolios fixed, operating profits are increasing and concave in the firm's own aggregate attribute. The return to entry on one line therefore depends on the firm's choices on its other lines, so portfolio probabilities generally do not factor into independent line probabilities. Heterogeneous attributes also make profits depend on portfolio composition rather than only on the number of active lines.

Expected profits and portfolio choice. Let $q_{mf}(a_f)$ denote firm f's portfolio probability in market m, and write $q_m = (q_{m1}, q_{m2}, q_{m3})$. For firm f, whose rivals are g and h, expected operating profits are

$$\bar{\pi}_{mf}(a_f; q_m) = \sum_{a_g \in \mathcal{A}_L} \sum_{a_h \in \mathcal{A}_L} q_{mg}(a_g) q_{mh}(a_h)\, \pi_f(a_f, a_g, a_h). \tag{G.20}$$

Under the portfolio-level extreme-value assumption, the implied choice probability is

$$P_{mf}(a_f; \theta, q_m) = \frac{\exp\{\bar{\pi}_{mf}(a_f; q_m) - C_{mf}(a_f; \theta)\}}{\sum_{\widetilde{a}_f \in \mathcal{A}_L} \exp\{\bar{\pi}_{mf}(\widetilde{a}_f; q_m) - C_{mf}(\widetilde{a}_f; \theta)\}}. \tag{G.21}$$

Equilibrium requires $q_{mf}(a_f) = P_{mf}(a_f; \theta, q_m)$ for every firm and portfolio. Conditioning on all other own-line choices gives

$$p_{mf\ell}(a \mid a_{f,-\ell}; \theta, q_m) = \frac{\exp\{\bar{\pi}_{mf}((a, a_{f,-\ell}); q_m) - C_{mf}((a, a_{f,-\ell}); \theta)\}}{\sum_{b=0}^{2} \exp\{\bar{\pi}_{mf}((b, a_{f,-\ell}); q_m) - C_{mf}((b, a_{f,-\ell}); \theta)\}}. \tag{G.22}$$

The full-portfolio normalizer cancels, leaving only three alternatives in the denominator.

Data construction and estimation. For each replication, I draw the entry-cost shifters and solve the equilibrium at θ_0, evaluating expected profits by summing over all rival portfolio pairs. Starting from uniform beliefs, I use a fixed-point iteration with one-half damping until the maximum discrepancy between beliefs and response probabilities is below 10^{-9}. I then draw portfolio-specific extreme-value shocks and record each firm's optimal portfolio. Both estimators use the same choices and observed primitives.

Both estimators alternate between parameter estimation under fixed rival beliefs and updates of those beliefs. The full-portfolio estimator maximizes the joint portfolio likelihood and evaluates expected profits and response probabilities by enumeration. The conditional estimator maximizes the sum of observed line-conditional log probabilities from equation (G.22), using sampled rival portfolios to approximate expected profits.

For the conditional estimator, each iteration uses $S = 64$ joint rival draws. The same draws are retained across candidate own portfolios and throughout the parameter optimization. Given the updated parameters and fixed rival beliefs, a separate Gibbs chain for each firm and market updates one line at a time using equation (G.22). Other own-line choices take their current sampled values. A firm's chain does not condition on another firm's latest sampled portfolio.

I discard 100 burn-in sweeps and retain one complete portfolio per sweep for $T = 64$ sweeps. The resulting empirical portfolio distributions update beliefs with one-half damping. Complete portfolios preserve dependence across a firm's lines. Expected profits are computed only for portfolios needed by the likelihood or Gibbs updates and are reused within the iteration. Beliefs are stored as a uniform component and weighted sampled portfolios, avoiding a dense vector over all 3^L alternatives.

Both estimators start from uniform beliefs. Convergence requires the maximum absolute parameter change to fall below 10^{-4} and the largest monitored conditional-probability change to fall below 5×10^{-3} for two consecutive iterations, with successful parameter optimization in both. Conditional probabilities are monitored at the observed portfolio and eight fixed random probe portfolios for each firm and market.

Monte Carlo results. Table A4 reports average parameter estimates, bias, root mean squared error, and computational performance. The two estimators produce similar estimates and RMSEs across the three specifications. Average estimates are generally close to the true values, although individual-replication errors remain appreciable for some parameters, particularly β_z when $L = 3$. All reported runs satisfy the numerical stopping rule.

The conditional estimator is slower at these portfolio sizes, but the runtime gap narrows substantially as the choice set expands. As L increases from three to five, average estimation time rises from 0.55 to 165.47 seconds for full-portfolio estimation and from 151.32 to 1,037.20 seconds for conditional estimation with Gibbs sampling. These times include profit calculations, parameter estimation, and belief updates, excluding data generation and diagnostics.

Enumeration benefits from the inexpensive profit function, reuse of operating profits across markets and iterations, and vectorized expectation calculations. Gibbs sampling incurs additional costs from sequential line updates and candidate evaluations. Its potential computational advantage comes from avoiding the exponential growth of the full portfolio set: with K actions on each of L lines, enumeration evaluates K^L own portfolios, whereas one Gibbs sweep compares KL candidate alternatives. Multiple sweeps are required, but none requires enumerating the full choice set.

Table A4: Monte Carlo estimates and computational cost

	$L = 3$		$L = 4$		$L = 5$	
	Full	Gibbs	Full	Gibbs	Full	Gibbs
Panel A. Parameter recovery						
κ ($\kappa_0 = 1.10$)	1.090	1.090	1.077	1.077	1.100	1.100
Bias	-0.010	-0.010	-0.023	-0.023	0.000	0.000
RMSE	0.031	0.031	0.045	0.045	0.036	0.036
β_x ($\beta_{x0} = 0.45$)	0.449	0.455	0.499	0.496	0.431	0.429
Bias	-0.001	0.005	0.049	0.046	-0.019	-0.021
RMSE	0.137	0.137	0.069	0.073	0.069	0.069
δ_2 ($\delta_{20} = 0.35$)	0.330	0.330	0.344	0.345	0.334	0.334
Bias	-0.020	-0.020	-0.006	-0.005	-0.016	-0.016
RMSE	0.055	0.056	0.029	0.029	0.050	0.049
Panel B. Computation						
Convergence rate (%)	100	100	100	100	100	100
Average estimation time (seconds)	0.55	151.32	5.47	403.47	165.47	1037.20
Time ratio (Gibbs/Full)	276.07		73.70		6.27	
Average NPL iterations	13.0	16.0	12.0	16.9	10.0	13.0

Notes: Each replication contains 500 markets and three firms, with three actions per line. Full and Gibbs denote full-portfolio NPL and conditional NPL with Gibbs sampling. Estimates, times, and iterations are averages over the same 100 replications for $L = 3, 4, 5$; the latter results are preliminary. Bias and RMSE use the stated true values. Gibbs uses 64 rival draws, 100 burn-in sweeps, and 64 retained sweeps per iteration. Convergence refers to the numerical stopping rule. Estimation time excludes data generation and diagnostics.

Robustness: random-coefficient demand. I also replace the baseline profit function with profits generated by random-coefficient logit demand, using PyBLP to compute demand and pricing outcomes (Conlon and Gortmaker, 2020). The entry-cost specification and portfolio shocks remain unchanged. Demand parameters and marginal costs are known to the economist; only entry-cost parameters are estimated.

For the products offered under portfolio configuration a, consumer utility is

$$u_{ij} = v_j - \alpha_i p_j + \epsilon_{ij}, \qquad v_j = 2 + \log x_j, \qquad \alpha_i = 1 + \sigma_\alpha \nu_i, \qquad \nu_i \sim U[-1, 1], \tag{G.23}$$

where $\sigma_\alpha \in [0, 1)$, the outside option has deterministic utility zero, and ϵ_{ij} is independently standard type-I extreme value. Marginal cost is normalized to $c_j = 1$. Given each portfolio configuration, firms set prices under multiproduct Bertrand competition. At the selected numerical pricing equilibrium $p^*(a)$, operating profits are

$$\pi_f(a) = D \sum_{j \in \mathcal{J}_f(a_f)} \Big(p_j^*(a) - c_j \Big) \, s_j(p^*(a), a), \qquad D = 20. \tag{G.24}$$

Consumer heterogeneity is integrated using a common fixed quadrature rule, with accuracy checked against a denser rule. Data generation and both estimators use the same numerical profit function. Profits are cached and reused across markets and entry-cost parameter evaluations.

Table A5 shows comparable parameter recovery across the two estimators, using 50 replications for both $L = 3$ and $L = 4$. Each previously unevaluated portfolio configuration requires solving a pricing equilibrium, making both data generation and estimation costly. I therefore restrict this exercise to $L = 3, 4$. At these portfolio sizes, Gibbs sampling visits much of the configuration space, limiting savings in pricing calculations while adding sequential sampling and belief-update costs.

G.3 Estimation Sample and Computational Implementation

This subsection describes the implementation for the washing-machine entry model. The portfolio payoff and line-conditional probabilities are defined in Section 6.2; the computational procedure uses those expressions directly.

Estimation sample. The entry risk set consists of potential firm–capacity–spin-speed lines observed over the estimation period. Each eligible line-quarter permits no entry or the introduction of one product in class A, B, C, or D. Line-quarters with multiple observed introductions cannot be represented by this action space and are excluded from the likelihood and Gibbs updates. Their actual products remain in the background portfolio, so they continue to affect demand and profits.

Table A5: Monte Carlo estimates and computational cost: Random-coefficient demand

	$L = 3$		$L = 4$	
	Full	Gibbs	Full	Gibbs
Panel A. Parameter recovery				
κ ($\kappa_0 = 1.10$)	1.084	1.084	1.079	1.079
Bias	-0.016	-0.016	-0.021	-0.021
RMSE	0.069	0.070	0.068	0.067
β_z ($\beta_{z0} = 0.45$)	0.446	0.450	0.448	0.449
Bias	-0.004	0.000	-0.002	-0.001
RMSE	0.105	0.104	0.082	0.083
δ_2 ($\delta_{20} = 0.35$)	0.343	0.344	0.348	0.347
Bias	-0.007	-0.006	-0.002	-0.003
RMSE	0.057	0.058	0.039	0.039
Panel B. Computation				
Paired replications	50	50	50	50
Convergence rate (%)	100	100	100	100
Average estimation time (seconds)	284.61	429.22	6343.81	13180.82
Time ratio (Gibbs/Full)	1.51		2.08	
Average NPL iterations	14.0	13.4	13.6	12.6

Notes: Each replication contains 500 markets and three firms, with three actions per line. Operating profits arise from random-coefficient logit demand with endogenous Bertrand prices computed using PyBLP. Full and Gibbs denote full-portfolio NPL and conditional NPL with Gibbs sampling. Both columns use 50 replications for $L = 3, 4$. Estimates, times, and iterations are replication averages. Bias and RMSE use the stated true values. Both estimators use 64 rival draws; Gibbs uses 100 burn-in sweeps and 64 retained sweeps per iteration. Starting values vary across replications and are shared across estimators within each replication.

Other eligible lines of the same firm remain in estimation. The reported likelihood contains 469 line-quarter observations from a broader risk set of 510 records.

Predetermined product attributes and inherited designs are held fixed when evaluating alternative entry actions. First introductions of an energy class permit redesign as specified in the model; repeated introductions inherit the existing class-specific design. The neural network described in Appendix F.1 supplies continuation payoffs net of design costs, so those costs are not subtracted again at the entry stage.

Expected payoffs and conditional likelihood. For each candidate own portfolio, I average continuation payoffs over complete rival-portfolio draws. The expected monetary payoff is multiplied by β_π, and entry costs are subtracted in the normalized units of equation (16). Payoffs are averaged before applying the logit transformation. The same rival draws are used across alternative own portfolios within each iteration.

The likelihood uses the line-conditional probabilities in equation (25), holding the firm's other line choices at their observed values. Their product is a composite conditional likelihood rather than a joint portfolio likelihood. Given rival beliefs, the payoff index is affine in β_π and the entry-cost

coefficients. I maximize the constrained criterion using three starting values and impose the cost restrictions specified in the model.

Gibbs sampling and belief updates. Given the current parameters and rival beliefs, each firm's Gibbs chain visits its eligible lines in random order. At each update, it draws an action from the line-conditional distribution given the current sampled choices on the firm's other lines. A complete portfolio is recorded after each sweep. The procedure discards 100 burn-in sweeps, retains 64 portfolios without thinning, and uses 64 joint rival draws to integrate expected payoffs.

Sampling complete portfolios preserves dependence across a firm's lines. Independence across firms is the maintained mixed-strategy assumption. Rival beliefs remain fixed during a firm's sampling step; the chain does not respond to another firm's latest sampled action. After all firms have been updated, the sampled portfolio distributions determine rival beliefs for the next estimation iteration.

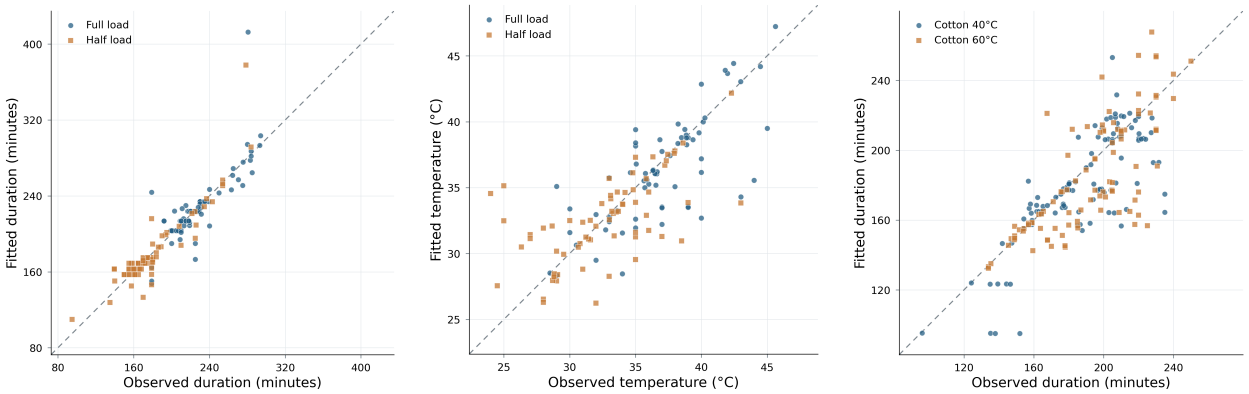
The algorithm alternates between parameter estimation and belief updates until relative parameter change is below 0.01 and the largest monitored conditional-probability change is below 0.02 for two consecutive iterations, subject to a maximum of 20 iterations. Monitoring covers the observed own portfolios and fixed probe configurations. An additional update with an independent random seed checks the sensitivity of monitored probabilities to simulation draws. These checks assess numerical stability over the monitored configurations rather than establishing full-chain mixing.

Conditional inference. Entry-cost standard errors use a composite-likelihood sandwich covariance, with the Hessian and score contributions projected onto the tangent space of binding cost restrictions. The nonclustered calculation uses line-quarter score contributions; a firm-clustered alternative aggregates them within each of the seven firms. A delta-method transformation converts normalized entry costs into monetary units.

These calculations condition on fitted beliefs, the selected active constraints, the continuation-payoff predictor, and upstream estimates. They do not incorporate uncertainty from equilibrium-belief feedback or active-set selection. They therefore differ from the full equilibrium-adjusted inference described in Theorem 3. With only seven firm clusters, the clustered estimates provide a sensitivity comparison rather than a definitive small-sample correction.

G.4 Model Fit

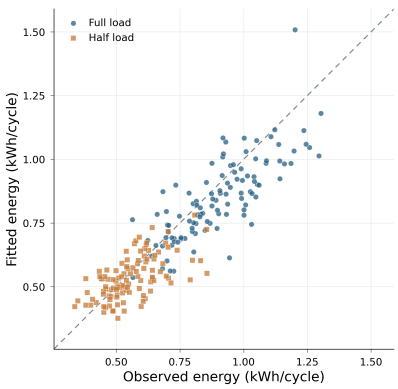
Figure G6: Observed and Fitted Program Inputs and Energy Consumption



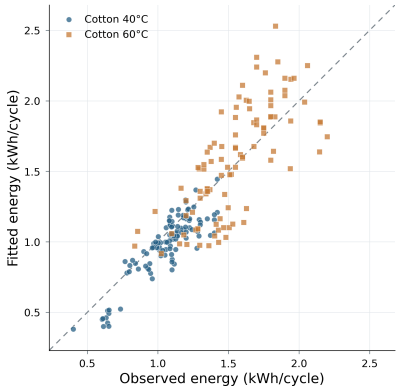
(a) Eco duration

(b) Eco temperature

(c) Cotton duration



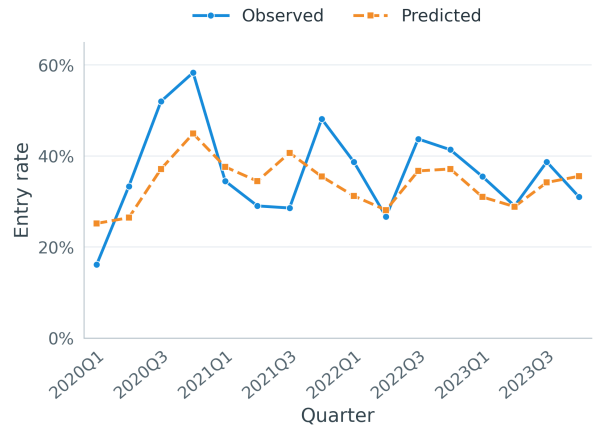
(d) Eco energy consumption



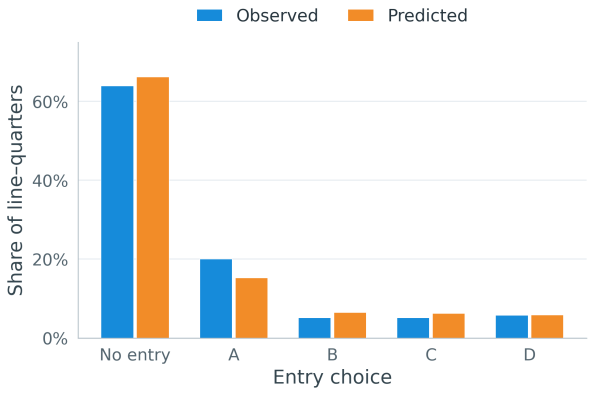
(e) Cotton energy consumption

Notes: Horizontal axes show observed values; vertical axes show model-implied values for product introductions and A^{+++} baseline products. Colors and markers distinguish full and half loads for Eco and Cotton 40°C and Cotton 60°C for Cotton. Duration is measured in minutes, temperature in degrees Celsius, and energy consumption in kWh per cycle. Dashed lines denote equality. Energy consumption is a fit diagnostic and does not enter the estimation criterion.

Figure G7: Entry Model Fit



(a) Quarterly entry rates



(b) Entry-choice shares



(c) Observed choices by reference class



(d) Predicted choices by reference class

Notes: Panels (a) and (b) use 469 eligible line–quarters pooled across the transition and post-policy periods. Quarterly rates use the eligible observations in each quarter. Predicted shares average line-conditional choice probabilities evaluated at observed own other-line actions, integrating over simulated rival portfolios. Panels (c) and (d) use the 429 observations with a single reference class; each column reports the distribution of current entry choices conditional on that reference class. Predicted probabilities are not unconditional frequencies from simulated entry histories.

Table A6: Selected Design Moments: Observed and Fitted

Moment		A^{+++}	A	B	C	D
Cotton 40°C duration (min)	Observed	176.46	191.20	194.75	199.48	183.94
	Fitted	178.76	198.48	195.59	196.81	171.96
Cotton 60°C duration (min)	Observed	177.11	183.48	190.39	199.29	190.47
	Fitted	179.33	186.26	183.62	195.03	184.26
Eco duration (min)	Observed	219.80	191.20	190.05	190.21	188.74
	Fitted	218.91	183.45	185.61	191.45	187.76
Eco temperature (°C)	Observed	35.75	31.46	35.43	36.40	38.50
	Fitted	35.84	32.09	34.19	34.75	34.47
Eco energy (kWh/cycle)	Observed	0.741	0.482	0.546	0.630	0.723
	Fitted	0.742	0.472	0.523	0.581	0.642
Cotton 40°C energy (kWh/cycle)	Observed	0.991	1.120	1.059	1.133	0.998
	Fitted	0.953	1.097	1.023	1.032	0.917
Cotton 60°C energy (kWh/cycle)	Observed	1.421	1.674	1.571	1.603	1.554
	Fitted	1.527	1.738	1.584	1.653	1.517

Notes: Observed and fitted values are data and model-implied means, respectively. Conditional energy efficiency is held fixed. A^{+++} means use baseline products; new-label means weight first class introductions equally. New-label temperature means exclude initial transitions from A^{+++}. Eco inputs average full and half loads; Eco energy uses the same loads for A^{+++} and regulatory full-, half-, and quarter-load weights for new-label classes.

H Counterfactual Setup and Computation

This appendix describes the policy experiments and their numerical implementation for 2020Q1–2023Q4. Demand, engineering, design-cost, and entry-cost parameters are held fixed across experiments, subject to the calibration qualification below. Firms respond through product introductions, adjustable program inputs, and prices; households respond through purchases and program use. All policy changes are measured relative to the model-implied current-policy baseline.

H.1 Policy experiments

I consider five counterfactual environments. The first removes the Eco-duration cap while retaining labels and efficiency requirements. The second removes label utility and efficiency-index restrictions while retaining the duration cap. The third broadens the efficiency index to include Cotton programs and removes the duration cap. The fourth replaces labels, efficiency restrictions, and the duration cap with an energy-based producer tax. The fifth removes the 2021 reform, retaining the pre-reform labeling and regulatory environment with program designs frozen at their reference values. The last experiment therefore differs from the other counterfactuals, which allow the program adjustments supported by the estimated design model.

Under the weighted-index policy,

$$EEI_j(w) = \frac{100\left(w_E e_{jE} + w_{40} e_{j40} + w_{60} e_{j60}\right)}{S(k_j)}, \qquad (w_E, w_{40}, w_{60}) = (0.80, 0.10, 0.10), \tag{H.1}$$

where $S(k_j)$ is the capacity-dependent reference consumption. Eco energy uses the regulatory load weights, while Cotton energy is measured at full load. Class thresholds remain unchanged. These weights define the reported experiment and are not estimated optimal weights.

The producer tax adds a levy to marginal cost:

$$\mathcal{T}_j = \tau_E \omega_E e_{jE} + \tau_C \left(\omega_{40} e_{j40} + \omega_{60} e_{j60}\right), \tag{H.2}$$

$$mc_j^\tau = mc_j + \mathcal{T}_j. \tag{H.3}$$

The reported experiment sets $\tau_E = \tau_C = 50$, with equal program weights $\omega_E = \omega_{40} = \omega_{60} = 1/3$. The levy is paid once per machine sold and is measured in euros per weighted kWh of per-cycle consumption; it is not a tax per tonne of carbon dioxide. It is added after the multiplicative production-cost shock is realized, and revenue is rebated to households through lump-sum transfers.

The weighted-index and tax experiments change several regulatory instruments simultaneously. Their comparisons with current policy therefore measure the effects of policy packages. Within the tax regime, zero tax yields the highest consumer welfare net of electricity expenditure among

the rates evaluated. This comparison holds the removal of labels, efficiency restrictions, and the duration cap fixed.

H.2 Market states and product outcomes

Each quarterly experiment conditions on the observed history used in entry estimation. This history determines incumbent products, entry eligibility, previous class introductions, and reference designs. Current introductions are endogenous, but simulated introductions do not replace observed histories in subsequent quarters. The results therefore describe equilibrium responses conditional on the observed sequence of market states, rather than a simulated industry transition.

Each eligible product line permits no entry or introduction in class A, B, C, or D. Line–quarters excluded from entry estimation remain fixed background products. When labels and efficiency restrictions are removed, class identifiers continue to distinguish candidate products and their entry costs, but no longer generate label utility or impose class-specific efficiency requirements.

Conditional on the portfolio, adjustable inputs include Eco temperature and duration and Cotton 40°C and Cotton 60°C duration. Conditional energy efficiency, water use, and other fixed inputs remain unchanged. The estimation sample's restrictions on Cotton 60°C duration also remain in place. Repeated class introductions inherit existing designs according to the reference rules in Appendix F.2. Capacity is fixed at the candidate-line level, so changes in average capacity reflect the composition of introductions.

Policy-specific neural networks approximate the design equilibrium and continuation payoffs using the architecture described in Appendix F.1. Continuation payoffs are already net of design costs, which are not deducted again when evaluating entry. For the portfolio draws used in outcome accounting, I recover predicted designs and solve the Bertrand pricing equilibrium before calculating purchases, program use, electricity consumption, and operating profits. Eco and Cotton energy are evaluated using their respective engineering mappings; changes in Eco inputs do not mechanically alter Cotton conditional efficiency.

H.3 Entry equilibrium and outcome integration

Entry equilibria use the portfolio payoffs and line-level conditional probabilities defined in equations (23) and (25). The eight estimated entry parameters are held fixed: the common launch cost, the profit coefficient, and the six switching premiums. Monetary continuation payoffs are multiplied by β_π, and portfolio-shock scale remains normalized to one.

At each belief iteration, I approximate expected payoffs using 64 rival-portfolio draws. Each draw samples a complete portfolio from each rival's current distribution, preserving dependence

among that rival's line choices. Holding these draws fixed, I update each firm's portfolio through random-order Gibbs sweeps. After 100 burn-in sweeps, I retain 64 portfolios, one per sweep, to form its updated empirical distribution. All firms respond to the incoming rival beliefs, and parameters are not reestimated. The Gibbs procedure follows Appendix G.3.

Beliefs are initialized using the observed entry rate, with initial entry probability divided equally across classes. I monitor conditional probabilities at the observed portfolio, the all-no-entry portfolio, and two fixed random portfolios. The largest change must be below 0.02 for two successive iterations, with a maximum of 30 iterations. I then repeat the final update using an independent seed and require both conditional-probability and marginal-probability gaps to remain below 0.02. Common random-number streams across belief iterations reduce simulation variation.

The 64 policy–quarter solves for the first four experiments pass these checks. The largest final iteration gap is 0.00013; the largest independent-seed conditional and marginal gaps are 0.00380 and 0.00493, respectively. These diagnostics assess numerical stability at the monitored portfolios rather than equilibrium uniqueness or complete mixing.

Outcome accounting uses 16 retained joint portfolio draws per policy and quarter. Each draw combines selected entrants with the fixed background portfolio. Outcomes integrate over household heterogeneity and production-cost shocks, are averaged across portfolio draws, and are then summed over quarters. The draws represent entry uncertainty within each observed market state, not independent industry histories.

Entry convergence and design accuracy are separate numerical considerations. Solving prices at predicted inputs does not establish that those inputs form an exact constrained design equilibrium.

I Welfare Accounting

I report consumer welfare net of electricity expenditure, discounted electricity expenditure, and operating profits. Welfare adjustments evaluate the purchase and program choices generated by the estimated behavioral model; they do not change those choices.

I.1 Purchase surplus and the label-energy correction

Let $\bar{u}^s_{ijmt}$ denote deterministic purchase utility under policy s, including the program inclusive value and any active label utility. With the purchase outside option normalized to zero and α_i denoting

the marginal utility of one euro, gross purchase surplus is

$$CS^s_{\text{gross}} = \sum_{t,m} M_{mt} \mathbb{E}_s \left[\int \frac{1}{\alpha_i} \log \left(1 + \sum_{j \in \mathcal{J}^s_{mt}} \exp\{\bar{u}^s_{ijmt}\} \right) dF_m(i) \right]. \tag{I.1}$$

Here, $\mathbb{E}_s$ integrates over simulated portfolios and market shocks. The estimated price coefficient is converted to euro units before applying this expression.

To separate perceived electricity savings from other label value, I project the estimated new-label coefficients on class-level Eco energy savings:

$$\widehat{\beta}_c = \gamma x_c + r_c, \qquad x_c = \bar{e}^E_D - \bar{e}^E_c, \qquad c \in \{A, B, C\}. \tag{I.2}$$

Energy is measured per loaded kilogram. Class means use the 2021Q1–2023Q4 demand sample after collapsing regional duplicates and restricting attention to capacity–quarter cells containing all four classes. Equal cell weights hold capacity and quarter composition constant across classes. The main correction uses ordinary least squares without an intercept; a covariance-weighted projection provides a sensitivity calculation.

I interpret $\widehat{\gamma} x_c$ as perceived energy-saving utility and retain r_c as nonpecuniary label value. This is an accounting assumption: the projection cannot separately identify financial savings, environmental preferences, and other attributes associated with labels. The projection is held fixed across policies, with $x_D = 0$ and no truncation of residuals.

Let P^s_{ijmt} denote the original purchase probability and λ^s_{jmt} indicate whether new-label utility is active. The monetary correction is

$$B^s = \sum_{t,m} M_{mt} \mathbb{E}_s \left[\int \sum_{j \in \mathcal{J}^s_{mt}} P^s_{ijmt} \frac{\lambda^s_{jmt} \widehat{\gamma} x_{c(j)}}{\alpha_i} \, dF_m(i) \right]. \tag{I.3}$$

It is zero when new-label utility is absent. Subtracting this probability-weighted term preserves the original choices and their associated taste-shock selection.

Consumer welfare before electricity expenditure is

$$CW^s_{\text{preE}} = CS^s_{\text{gross}} - B^s + R^s, \tag{I.4}$$

where R^s is the tax rebate. The program inclusive value remains in gross surplus, so no separate lifetime washing-benefit term is added. This accounting treats retained program utility and residual label value as nonpecuniary benefits; it does not remove an additional electricity-cost component from program fixed effects.

I.2 Electricity consumption and expenditure

Electricity accounting covers Eco, Cotton 20°C, Cotton 40°C, Cotton 60°C, and Easy Care. Let $\mathcal{R}$ contain these programs, $\bar{N}$ denote mean annual washing frequency, and p^s_{ijrmt} denote program-choice probabilities conditional on purchasing product j. Annualized buyer electricity is

$$E^s_{\text{buy}} = \sum_{t,m} M_{mt} \mathbb{E}_s \left[\int \sum_{j \in \mathcal{J}^s_{mt}} P^s_{ijmt} \bar{N} \sum_{r \in \mathcal{R}} p^s_{ijrmt} e^s_{jrmt} \, dF_m(i) \right]. \tag{I.5}$$

Program probabilities are not renormalized: electricity used by programs outside $\mathcal{R}$ is outside the reported coverage.

Machines retained by nonbuyers. Nonbuyers use a common reference basket constructed from the observed 2021Q1 portfolio, weighted by unit sales. Its annual electricity consumption, $\bar{E}^{\text{ref}}$, is calculated using survey mean annual wash counts for each covered program. The basket is fixed across policies and serves as a proxy for retained machines rather than an observed household appliance stock.

If Q^s is total purchases across the accounting observations, total annualized electricity consumption is

$$E^s = E^s_{\text{buy}} + \left(\sum_{t,m} M_{mt} - Q^s \right) \bar{E}^{\text{ref}}. \tag{I.6}$$

Thus, retained-machine electricity changes only through the number of nonbuyers. The totals sum annualized use associated with quarterly decision populations; they do not represent electricity use in a single calendar year or distinct households over the sample period.

Electricity is valued at 0.30 euros per kWh over a 12-year service life, discounted at 5 percent:

$$C^s_E = 0.30 \, A_{12} \, E^s, \qquad A_{12} = \sum_{y=1}^{12} (1.05)^{-y}. \tag{I.7}$$

Washing frequency, the electricity price, service life, and discount rate are common across policies. Reported consumer welfare is

$$CW^s_{\text{net}} = CS^s_{\text{gross}} - B^s + R^s - C^s_E, \tag{I.8}$$

$$\Delta CW_{\text{net}} = \Delta CS_{\text{gross}} - \Delta B + \Delta R - \Delta C_E. \tag{I.9}$$

Electricity expenditure is therefore already included in net consumer welfare and should not be subtracted again.

I.3 Profits and reported comparisons

Operating profits are sales revenue less variable production costs and the producer tax:

$$\Pi^s_{\text{op}} = \sum_{t,m} \mathbb{E}_s \left[\sum_{j \in \mathcal{J}^s_{mt}} n^s_{jmt} \left(p^s_{jmt} - mc^s_{jmt} - \mathcal{T}^s_{jmt} \right) \right], \tag{I.10}$$

where mc^s_{jmt} excludes the tax. These profits do not deduct entry or design costs and therefore differ from the continuation payoffs used in entry decisions.

Tax revenue equals the aggregate rebate,

$$R^s = \sum_{t,m} \mathbb{E}_s \left[\sum_j n^s_{jmt} \mathcal{T}^s_{jmt} \right].$$

The levy is deducted once from firms' profits and returned once to households. Adding consumer welfare and operating profits cancels the transfer but does not yield comprehensive social welfare, because entry and design costs, environmental damages, and electricity-supply surplus are omitted.

The current-policy baseline is evaluated using the same model and accounting conventions as the counterfactuals. Product-attribute statistics are unweighted means over eligible first firm–line–class introductions, pooled across quarters and portfolio draws. Repeat introductions are excluded. Eco temperature and duration refer to full load; Eco energy uses regulatory load weights, and Cotton energy is measured at full load. These statistics describe introductions rather than the installed stock.

Table A7 reports baseline levels and signed changes under the five counterfactuals. Panel A covers product attributes and energy per cycle. Panel B reports net consumer welfare, electricity expenditure, and operating profits in millions of euros. For outcome Y, level changes are $Y^s - Y^0$ and percentage changes are

$$100 \, \frac{Y^s - Y^0}{|Y^0|}.$$

Negative expenditure changes indicate savings. Percentage changes have outcome-specific denominators and cannot be combined directly. The comparisons reflect joint equilibrium responses rather than separate contributions from redesign, program use, purchases, and entry.

Table A7: Counterfactual Changes Relative to Model-Implied Current Policy

Outcome	Unit	Baseline	No duration cap	No labels or efficiency restrictions	Weighted EEI	Carbon tax	No 2021 reform
Panel A. Product attributes							
Eco temperature	°C	31.69	-1.81	+4.51	-6.23	+0.51	+5.89
Capacity	kg	8.08	-0.02	+0.04	+0.00	+0.03	-0.05
Eco duration	min	208.91	+51.93	+1.89	+38.88	+54.07	+27.61
Cotton 40°C duration	min	186.39	-5.61	-9.19	-59.39	-30.88	-13.58
Cotton 60°C duration	min	168.70	-3.98	+2.60	-8.13	-2.99	+2.19
Eco energy/cycle	kWh	0.558	-0.017	+0.090	-0.099	+0.024	+0.132
Cotton 40°C energy/cycle	kWh	0.986	-0.021	-0.059	-0.246	-0.118	+0.009
Cotton 60°C energy/cycle	kWh	1.470	-0.026	-0.026	-0.052	-0.003	-0.123
Panel B. Welfare and expenditure changes							
Electricity expenditure	€ million	6,474.90	-71.87	+33.10	-21.26	+88.04	-195.74
	% of absolute baseline	—	-1.11	+0.51	-0.33	+1.36	-3.02
Consumer welfare	€ million	-5,444.58	+63.24	+59.52	+76.59	-74.75	+231.13
	% of absolute baseline	—	+1.16	+1.09	+1.41	-1.37	+4.25
Operating profits	€ million	810.13	+52.74	-106.11	+40.92	-365.61	-100.39
	% of absolute baseline	—	+6.51	-13.10	+5.05	-45.13	-12.39

Notes: The baseline reports model-implied current-policy levels; policy columns report signed changes. Product attributes are unweighted means of first class introductions within lines. The no-reform case fixes the pre-reform portfolio and designs while reoptimizing prices. Eco inputs refer to full load; Eco energy uses the applicable load weights. Consumer welfare is net of electricity expenditure and includes the label-energy correction and tax rebates. Electricity covers buyers and retained machines.

References

Besag, J. (1975). Statistical analysis of non-lattice data. *The Statistician*, 24(3):179–195.

Brook, D. (1964). On the distinction between the conditional probability and the joint probability approaches in the specification of nearest-neighbour systems. *Biometrika*, 51(3/4):481–483.

Casella, G. and George, E. I. (1992). Explaining the Gibbs sampler. *The American Statistician*, 46(3):167–174.

Kraft, D. (1988). A software package for sequential quadratic programming. Technical Report DFVLR-FB 88-28, Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt, Institut für Dynamik der Flugsysteme, Oberpfaffenhofen, Germany.

van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press.

Van Holsteijn en Kemna, Vlaamse Instelling voor Technologisch Onderzoek, Viegand Maagøe, and Wuppertal Institut (2014). "Omnibus" Review Study on Cold Appliances, Washing Machines, Dishwashers, Washer-Driers, Lighting, Set-top Boxes and Pumps. Final report, European Commission, Directorate-General for Energy, Brussels and Delft. 12 March 2014.

Varin, C., Reid, N., and Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica*, 21(1):5–42.

Zaheer, M., Kottur, S., Ravanbakhsh, S., Póczos, B., Salakhutdinov, R., and Smola, A. J. (2017). Deep sets. In *Advances in Neural Information Processing Systems*, volume 30.