Online Appendix

Appendix A. Comparative Statics Results

Information Acquisition Costs. Naturally, we should expect that consumers will collect and process more energy information the lower the costs to do so. In particular, a consumer should always choose to be fully informed if there are no extra costs. The present model is consistent with this intuition, whether ENERGY STAR information is available or not.

Proposition 1. (i) *Suppose that ENERGY STAR information is not available. If* $\mathcal{K}(e = U) = \mathcal{K}(e = I)$*, it is optimal for the consumer to select* $e = h$.

(ii) *Suppose that ENERGY STAR information is available. If* $\mathcal{K}(e = U) = \mathcal{K}(e = ES) = \mathcal{K}(e = I)$*, it is optimal for the consumer to select* $e = I$*. Moreover, if* $\mathcal{K}(e = U) = \mathcal{K}(e = ES)$*,* $e = ES$ *is strictly better than* $e = U$ *for the consumer.*

Proof. This is true by the fact that the expectation of the maximum of random variables is always greater that the maximum of their expectations. In particular, consider some set of random variables $\{Y_1, Y_2, ..., Y_k\}$. The distribution of $\max_{1 \le j \le k}\{Y_j\}$ (first order) stochastically dominates the distribution of Y_l for any $l \in \{1, ..., k\}$. This implies that $E[\max_{1 \le j \le k} Y_j] \ge E[Y_l]$ for $l = 1, ..., k$, and thus,

$$E[\max_{1 \le j \le k} Y_j] \ge \max_{1 \le j \le k} E[Y_j]. \tag{14}$$

To show (i), it suffices to show that

$$E_{\epsilon,\mathcal{C}}\left[\max_j\{U_{ij}(\delta_j, \eta P_j, C_j, \epsilon_{ij})\}\right] \ge E_\epsilon\left[\max_j\{E_{\mathcal{C}}\left[U_{ij}(\delta_j, \eta P_j, C_j, \epsilon_{ij})\right]\}\right],$$

which implies that (i) holds for $K = 0$. I show a stronger inequality; in particular, that for any ϵ_{ij},

$$E_{\mathcal{C}}\left[\max_j\{U_{ij}(\delta_j, \eta P_j, C_j, \epsilon_{ij})\}\right] \ge \left[\max_j\{E_{\mathcal{C}}\left[U_{ij}(\delta_j, \eta P_j, C_j, \epsilon_{ij})\right]\}\right].$$

This follows from (14) if I set

$$Y_j \equiv U_{ij}(\delta_j, \eta P_j, C_j, \epsilon_{ij}).$$

This concludes the proof for (i).

The proof for (ii) is similar.

Crowding-Out Effect. A simple, but important implication to the above result is that if the costs of processing and collecting ENERGY STAR information are lower than the costs of searching for energy costs, some consumers may prefer to select the maximum level of effort than to not collect information at all, but could prefer a medium level of effort than a maximum one. Formally,

Corollary 1. *If $\mathcal{K}(e = ES) < \mathcal{K}(e = I)$, then for some consumers*

$$\mathcal{V}(e = U) < \mathcal{V}(e = I) < \mathcal{V}(e = ES)$$

Proof. The proof follows directly from Proposition 1.

This formally shows that the ENERGY STAR certification induces some consumers to be less informed and crowds out efforts to fully account for energy costs.

Uncertainty about Energy Costs. The present model stipulates that uncertainty in beliefs is the main driver that induces consumers to search for energy information. Therefore, the model should predict that the larger the uncertainty in beliefs, the more likely consumers are to search.

I now show this result. I focus on the case that consumers' beliefs become more uncertain, but remain unbiased. This scenario is notably consistent with **?**'s findings that shows that consumers' beliefs about future gas prices are on average unbiased, but largely uncertain.

Consider the following definition. Beliefs about X represented by a distribution $\hat{F}$ are more uncertain that beliefs F if F second order stochastically dominates $\hat{F}$:

$$\int_{\underline{x}}^{b} F(x)dx \geq \int_{\underline{x}}^{b} \hat{F}(x)dx \tag{15}$$

for all b.

Proposition 2. *If* $\hat{X}_j \sim \hat{F}, \forall j$ *and* $X_j \sim F, \forall j$, *with* $\int_{\underline{x}}^{b} F(x)dx \geq \int_{\underline{x}}^{b} \hat{F}(x)dx$ *for all* b, *then*

$$E_{\epsilon,\mathcal{X},S}\left[\max_j\{U_{ij}(\delta_j, \eta P_j, X_j(S), d, \epsilon_{ij})\}|\mathcal{I}(e)\right] \geq E_{\epsilon,\hat{\mathcal{X}},S}\left[\max_j\{U_{ij}(\delta_j, \eta P_j, \hat{X}_j(S), d, \epsilon_{ij})\}|\mathcal{I}(e)\right]$$

Proof. First note that by the definition of second order stochastically dominance, if X_j second-order stochastically dominates $\hat{X}_j$, and if X_j and $\hat{X}_j$ have the same mean, then $E[h(X_j)] \geq E[h(\hat{X}_j)]$ for all concave function h. Given that the maximum is a concave function, I then have $E[h(X_j, Y)] \geq E[h(\hat{X}_j, Y)]$ for any variable Y.

Proposition 2 simply says that the larger the variance in energy costs, the higher is the value of information. This also implies that ENERGY STAR will lead to more sub-optimal choices, in expectation, in choice sets where products are largely disperse in the energy efficiency characteristics space.

Appendix B. Data Cleaning and Manipulation

Creating a Random Sample. To perform the estimation of the demand model, a random sub-sample of the transactions is used. The sub-sample is constructed as follow.

First, the sub-sample is drawn from the set of transactions that fit the following criteria (the restricted sample):

- transactions made by consumers that are homeowners;
- transactions made by consumers living in single family housing units; and
- transactions made by consumers that made no more that one refrigerator purchase in any given year.

Second, the following stratified sampling method is used to create the sub-sample. For a given targeted sample size, I sample transactions for three different income groups:

- households with income of less than ≥$50,000;
- households with income between $50,000 and $100,000;
- and households with income of more than $100,000.

Average Electricity Prices. The use of average electricity prices is partly motivated by recent empirical evidence (**?**, **?**) that suggests that electricity consumers may in fact respond to variation in average prices, more than marginal prices. In the present case, the use of average electricity prices is also dictated by the fact that household's location is not perfectly known. Therefore, it is impossible to match households with their exact electricity tariff and infer marginal price.

Average electricity prices at the county level are computed as follow. Using form EIA-861 of the Energy Information Administration, I compute the average residential electric price for each electric utility operating in the US for the years 2008. I then match electric utility territories with each of the county where I sampled at least one store. For counties with only one electric utility, I use the average electricity price for this particular utility. For counties with several electric utilities, I take the arithmetic mean of each utility average price to construct the county level price.

Appendix C. Alternative Estimators

The Average Consumer. Adding an outside option to the conditional logit, I obtain a linear expression for the market shares in region r at time t (**?**):

$$ln(q_{jrt}) = \tau D_{jt} - \eta P_{jrt} + \psi R_{rt} X D_{jt} - \theta C_{jr} + \gamma_j + \alpha_r + \alpha_t + \zeta_{jrt} \quad (16)$$

where q_{jrt} is the quantity of refrigerator model j sold during week t in store r, and ζ_{jrt} is a market-specific unobservable. Not all refrigerator models sell every week in every store, the dependent variable thus takes the value zero for a large number of observations. Equation 16 is then estimated with a negative binomial model. The sample used for the estimation consists of all transactions observed during the period 2007-2009. Stores with a low number of sales were excluded, which left 545 stores for the estimation. Price time series for each refrigerator model are week and store specific. The choice set are store and month specific, and constructed using the same methodology than for the conditional logit model.

In the first specification, product, week, and store fixed effects are included, and the electricity prices averaged at the county level are used. The estimate of the price coefficient corresponds to an own-price elasticity of -2.89, which is about half the elasticity obtained with the conditional logit model. The present price elasticity has a different interpretation here given that an outside option has been added, and can be interpreted as being long-run. The size of the label effect is 0.041, and corresponds to a WTP for the ENERGY STAR label of 18$. This replicates closely the estimate obtained with the condition logit (19$). The effect of rebate is negative and not significant. The estimate of the coefficient on electricity costs is negative and significant, and implies a discount rate of 41%, which is lower than for the conditional logit model (62%).

The second specification replace the week fixed effects, with brand-week fixed effects. Like in the conditional logit model, it has few effects on the price coefficient, suggesting that marketing efforts are not an important source of bias. The coefficient for the ENERGY STAR label is larger under this specification.

The third specification uses electricity prices averaged at the state level. This impacts the coefficient on electricity costs. The coefficient still negative and significant, but implies a discount rate of 29%. Similar than in the condition logit model, measurement error in how electricity prices are measured has an important effect.

The fourth specification includes store-year fixed effects. This provides a better control for any region specific unobservables. Doing so has few impact on the coefficients.

Interaction with Demographics. Table 8 presents the results from a conditional logit model where the coefficient on prices is interacted with demographic information, in addition of the coefficients on electricity costs, and ENERGY STAR dummy. The results complement Table 4, and show that the main patterns hold.

TABLE 7. Negative Binomial Model

	(I)	(II)	(III)	(IV)
Price ($\hat{\eta}$)	-0.222***	-0.219***	-.219***	-0.219***
	(0.006)	(0.006)	(.006)	(0.006)
ENERGY STAR ($\hat{\tau}$)	0.041***	0.076***	.075***	0.077***
	(0.009)	(0.010)	(.010)	(0.010)
Rebate ($\hat{\psi}$)	-0.011	-0.012	-.010	-0.013
	(0.013)	(.013)	(.013)	(0.014)
Elect. Cost ($\hat{\theta}$)	-.532***	-0.533***	-.740***	-0.543***
	(0.115)	(.115)	(0.088)	(0.120)
Product FE	Yes	Yes	Yes	Yes
ZipCode FE	Yes	Yes	Yes	Yes
Week FE	Yes	No	No	No
BrandXWeek FE	No	Yes	Yes	Yes
Avg. Elec. Price	County	County	State	County
Observations	9.50e+06	9.50e+06	9.50e+06	9.50e+06
Nb Clusters	545	545	545	545
Interpretation: TO UPDATE				
Price Elasticity	-2.89	-2.85	-2.85	-2.85
WTP ES τ/η	18.0	34.7	34.2	35.2
Prob. Take Rebate ψ/η	-0.05	-0.05	-0.05	-0.06
Implied Discount Rebate	0.417	0.410	0.293	0.402

Note: † (p¡0.10) * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$)

TABLE 8. Conditional Logit: Interaction with Demographics

	Income <\$50,000			**Income** ≥\$50,000 & <\$100,000			**Income** ≥\$100,000		
Price	-0.586***			-0.496***			-0.404***		
	(0.022)			(0.013)			(0.012)		
Elect. Cost	0.445			0.222			1.643***		
	(0.320)			(0.287)			(0.361)		
Rebate	0.073***			0.020			-0.022		
	(0.021)			(0.016)			(0.018)		
ENERGY STAR	0.0179			-0.051			0.313***		
	(0.074)			(0.063)			(0.069)		
Interactions:	Elect. Cost	ENERGY STAR	Price	Elect. Cost	ENERGY STAR	Price	Elect. Cost	ENERGY STAR	Price
X Educ College	0.011	-0.066*	-0.003	-0.076	-0.020	0.001	-0.473***	-0.065**	0.008***
	(0.101)	(0.027)	(0.003)	(0.086)	(0.021)	(0.002)	(0.092)	(0.022)	(0.002)
X Educ Grad	-0.445	0.016	0.024***	-0.733***	-0.030	0.017***	-0.737***	-0.060*	0.012***
	(0.230)	(0.051)	(0.005)	(0.129)	(0.031)	(0.003)	(0.121)	(0.025)	(0.002)
X Inc +\$16.5K	-0.030	0.088*	0.015***	0.255*	0.037	0.010***	-0.308**	-0.059*	0.013***
	(0.138)	(0.036)	(0.004)	(0.124)	(0.030)	(0.003)	(0.106)	(0.024)	(0.002)
X Inc +\$33K	0.215	0.132***	0.011**	-0.072	0.055*	0.027***	-0.759***	-0.167***	0.023***
	(0.116)	(0.030)	(0.003)	(0.111)	(0.026)	(0.002)	(0.107)	(0.021)	(0.002)
X Age $_{>30,\leq 45}$	-0.350	-0.034	0.024***	-0.444**	-0.010	0.020***	-0.826***	-0.089	0.025***
	(0.197)	(0.056)	(0.006)	(0.161)	(0.044)	(0.004)	(0.201)	(0.051)	(0.004)
X Age $_{>45,\leq 55}$	-1.114***	-0.088	0.014*	-1.053***	-0.034	0.008*	-1.366***	-0.077	0.015***
	(0.198)	(0.058)	(0.006)	(0.164)	(0.043)	(0.004)	(0.208)	(0.052)	(0.004)
X Age $_{>55,\leq 70}$	-2.304***	0.042	0.013*	-2.592***	0.0010	0.013***	-2.520***	-0.043	0.008*
	(0.194)	(0.053)	(0.005)	(0.174)	(0.044)	(0.004)	(0.208)	(0.053)	(0.004)
X Age $_{>70}$	-3.392***	0.137*	-0.026***	-3.425***	0.164***	-0.015***	-3.127***	0.072	-0.021***
	(0.206)	(0.054)	(0.006)	(0.196)	(0.047)	(0.004)	(0.240)	(0.057)	(0.005)
X FamSize 2	0.342	0.109**	0.015**	0.331*	0.168***	0.012***	-0.588**	0.025	0.015***
	(0.177)	(0.042)	(0.005)	(0.139)	(0.032)	(0.003)	(0.179)	(0.039)	(0.003)
X FamSize 3-4	0.628***	0.074	0.026***	0.436**	0.140***	0.019***	-0.480**	-0.021	0.020***
	(0.168)	(0.041)	(0.004)	(0.136)	(0.032)	(0.003)	(0.175)	(0.037)	(0.003)
X FamSize 5 +	0.817***	0.060	0.033***	0.750***	0.112***	0.018***	-0.364*	-0.021	0.016***
	(0.183)	(0.046)	(0.005)	(0.152)	(0.033)	(0.003)	(0.183)	(0.038)	(0.003)
X Pol Dem.	0.080	-0.098*	-0.011*	0.224	-0.030	-0.012***	0.047	0.004	-0.009***
	(0.187)	(0.042)	(0.005)	(0.130)	(0.028)	(0.002)	(0.129)	(0.026)	(0.002)
X Pol Others	0.150	-0.106**	-0.013**	0.558***	-0.049	-0.011***	0.413**	-0.003	-0.009***
	(0.174)	(0.039)	(0.004)	(0.123)	(0.027)	(0.002)	(0.132)	(0.025)	(0.002)

Notes: Standard errors clustered at the store level. * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$)

Appendix D. FKRB's Estimator

The FKRB's estimator models the CDF of the random parameters as a mixture of point masses. Alternatively, FKRB propose to estimate a smooth density by modeling the distribution of the parameters as a mixture of normal densities. This requires to specify M normal basis functions with a predetermined mean and variance, and use simulation to construct the estimator.

In particular, the m^{th} basis function is defined as the product of the marginals of the K random parameters:

$$N(\beta|\mu^m, \sigma^m) = \prod_{k=1}^{K} N(\beta|\mu_k^m, \sigma_k^m) \tag{17}$$

The simulated choice probability is thus given by:

$$Q_{ijrt} \approx \sum_{m}^{M} \alpha_m \left(\frac{1}{S} \sum_{s=1}^{S} P^m_{ijrt|\beta^{m,s}} \right) \tag{18}$$

where $\beta^{m,s}$ is the s^{th} draw from the r^{th} normal basis. The estimate of α is thus given by:

$$\hat{\alpha} = argmin_{\alpha} \frac{1}{NJ} \sum_{i=1}^{I} \sum_{j=1}^{J} \left(y_{ijrt} - \sum_{m}^{M} \alpha_m \left(\frac{1}{S} \sum_{s=1}^{S} P^m_{ijrt|\beta^{m,s}} \right) \right) \tag{19}$$

s.t.

$$\sum_{m}^{M} \alpha_m = 1, \ \alpha_m \geq 0$$

For the present application, η, ψ and γ_j are taken as data, and they are fixed at the MLE estimates (Model 1, Table 2). The joint density of the parameters θ and τ is a mixture of $M = 108$ normal basis functions. To construct the basis functions, 9 marginals for the parameter θ, and 11 marginals for the parameter τ are used. The means of the marginals are defined relative to the MLE estimates of the coefficients $\hat{\tau}$ and $\hat{\theta}$. In particular, the means of the marginals for the parameter θ is the vector: $[2.25\hat{\theta}, 2\hat{\theta}, 1.75\hat{\theta}, 1.25\hat{\theta}, \hat{\theta}, 0.75\hat{\theta}, 0.5\hat{\theta}, 0.25\hat{\theta}, 0.05\hat{\theta}]$. The means of the marginals for the parameter τ is the vector: $[8\hat{\tau}, 6\hat{\tau}, 4\hat{\tau}, 2.5\hat{\tau}, 1.75\hat{\tau}, 1.25\hat{\tau}, \hat{\tau}, 0.75\hat{\tau}, 0.5\hat{\tau}, 0.25\hat{\tau}, 0]$. For each marginal, the standard deviation is set to a small value corresponding to 0.1% of the mean of the marginal. The estimation is carried with the matlab package lsqlin.

Appendix E. Policy Analysis

E.1. Emission Factors

Table 9. Emission Factors and Externality Costs

Non-baseload Output Emission Rates (U.S. Average)			
Pollutant	Estimate		Source
$CO2$	1,583 lb/MWh		
$CH4^a$	35.8 lb/GWh		
$N2O^a$	19.9 lb/GWh		USEPA, eGRID2007
$SO2$	6.13 lb/MWh		
NOx	2.21 lb/MWh		
Damage Cost (2008 $)			
Pollutant	Low Estimate	High Estimate	Source
$CO2$	$21.8/t	$67.1/t	**?**
$SO2$	$2,060/t	$6,700/t	low: **?**, high: USEPA[b]
NOx	$380/t	$4,591/t	low: **?**, high: DOE[c]

Notes: (a) Externality costs associated to $CH4$ and $N2O$ are assumed to be the same than for $CO2$. $CH4$ and $N2O$ are converted in CO2 equivalent using estimates of global warming potential (GWP). The GWP used for $CH4$ is 25, and the GWP used for $N2O$ is 298. Source: IPCC Fourth Assessment Report: Climate Change 2007. (b) Estimate used in the illustrative analysis of the 2012 regulatory impact analysis for the proposed standards for electric utility generating units. (c) Higher value of the estimate used in the Federal Rule for new minimum energy-efficiency standards for refrigerators (1904-AB79).

E.2. **Sensitivity Tests**

TABLE 10. The Opportunity Cost of Imperfect Energy Information: Low Electricity Costs, No Rebate

	Perfectly Informed vs Information Acquisition	**Perfectly Informed vs Uninformed**	**Perfectly Informed vs ENERGY STAR**
Consumer Surplus with Label Effect ($)			
Income <$50,000	13	18	5
≥$50,000 & <$100,000	8	12	12
≥$100,000	11	18	13
All Income	10	16	11
Consumer Surplus without Label Effect ($)			
Income <$50,000	20	24	32
≥$50,000 & <$100,000	9	12	20
≥$100,000	16	18	30
All Income	14	17	27
kWh/year purchased			
	-14	-21	-19
Externality Costs ($)			
	-11	-17	-15
Producer Surplus ($)			
	-21	-22	-47
Welfare ($)			
with label	4	15	-15
without label	8	16	1

Notes: In each column, a choice model is compared to a model where all consumers are perfectly informed. The first column compares the outcomes obtained with the information acquisition model. The second column compares a choice model where all consumers are uninformed. The third column compares a choice model where all consumers rely on ENERGY STAR. The table shows that the consumers surplus decreases in all cases when consumers are not perfectly informed, but profits increase. The externality costs increase under imperfect information.

Table 11. The Opportunity Cost of Imperfect Energy Information: High Electricity Costs, $50 Rebate

	Perfectly Informed vs Information Acquisition	**Perfectly Informed vs Uninformed**	**Perfectly Informed vs ENERGY STAR**
Consumer Surplus with Label Effect ($)			
Income <$50,000	52	63	41
≥$50,000 & <$100,000	31	44	39
≥$100,000	41	63	52
All Income	40	56	45
Consumer Surplus without Label Effect ($)			
Income <$50,000	60	67	72
≥$50,000 & <$100,000	31	42	49
≥$100,000	47	61	72
All Income	44	55	63
kWh/year purchased			
	-27	-37	-34
Externality Costs ($)			
	-11	-17	-15
Producer Surplus ($)			
	-30	-30	-59
Welfare ($)			
with label	21	42	0
without label	25	42	19

Notes: In each column, a choice model is compared to a model where all consumers are perfectly informed. The first column compares the outcomes obtained with the information acquisition model. The second column compares a choice model where all consumers are uninformed. The third column compares a choice model where all consumers rely on ENERGY STAR. The table shows that the consumers surplus decreases in all cases when consumers are not perfectly informed, but profits increase. The externality costs increase under imperfect information.

Table 12. The Opportunity Cost of Imperfect Energy Information: No Bunching at Previous Certification Requirement

	Perfectly Informed vs Information Acquisition	**Perfectly Informed vs Uninformed**	**Perfectly Informed vs ENERGY STAR**
Consumer Surplus with Label Effect ($)			
Income <$50,000	9	24	-20
≥$50,000 & <$100,000	6	15	3
≥$100,000	8	24	4
All Income	8	21	-2
Consumer Surplus without Label Effect ($)			
Income <$50,000	21	28	17
≥$50,000 & <$100,000	8	11	11
≥$100,000	14	19	20
All Income	14	18	16
kWh/year purchased			
	-15	-24	-15
Externality Costs ($)			
	-12	-19	-12
Producer Surplus($)			
	-20	-11	-41
Welfare ($)			
with label	-1	29	-32
without label	5	26	-14

Notes: In each column, a choice model is compared to a model where all consumers are perfectly informed. The first column compares the outcomes obtained with the information acquisition model. The second column compares a choice model where all consumers are uninformed. The third column compares a choice model where all consumers rely on ENERGY STAR. The table shows that the consumers surplus decreases in all cases when consumers are not perfectly informed, but profits increase. The externality costs increase under imperfect information.