Flexible customer willingness to pay for bundled smart home energy products and services

Ricardo A. Daziano

School of Civil and Environmental Engineering, Cornell University, Ithaca, NY 14853

Abstract

Energy markets are rapidly changing with smarter, connected, more reliable infrastructure and cleaner generation on the supply side, and more choice, greater control and enhanced flexibility for customers. This paper examines willingness to pay for bundled smart home energy products and information services, using data from a set of two discrete choice experiments that were part of a survey by the regional energy provider of upstate New York. To let the data reveal how preferences are distributed in the population, a logit-mixed logit model in willingness-to-pay space and a combination of observed and unobserved preference heterogeneity was specified and fitted. Results show that residents of Tompkins County are willing to pay more than in other counties for residential storage, and that for home energy management there is an important generational divide with millennials being much more likely to perceive the economic value in the smart energy technologies. The flexible logit-mixed logit estimates provide evidence of important heterogeneity in preferences: whereas most of the population has a positive –albeit rather low– valuation of smart energy products and services, there is a considerable percentage of customers with negative perceptions.

Keywords: discrete choice, seminonparametrics, willingness to pay, logit-mixed logit, smart energy packages, new energy markets
JEL classification: C11, D12, D47

Email addresses: `daziano@cornell.edu` (Ricardo A. Daziano), `Phone: +1 607 255-2018` (Ricardo A. Daziano)

1. Introduction

Energy markets are rapidly changing with smarter infrastructure and cleaner generation on the supply side, and more choice, greater control and enhanced flexibility for customers. Flexibility in the provision of home energy products and services is allowing utilities (energy service providers) to offer bundles that their residential and business customers can personalize. For example, taking advantage of the detailed information collected by smart meters, utilities are offering to their customers the possibility of checking their energy usage online at a granular level paired with personalized information that can be used to make more energy efficient choices. Customers are also now able to optimize the proportion of energy that is coming from renewable sources, as well as to make informed decisions regarding shifting energy use to off-peak times. In fact, smart energy management systems can automatically respond to dynamic pricing. Since residential customers have a plethora of options, energy service providers need to identify which smart home energy products and services best meet preferences and needs of their customer base.

This study focuses on customer willingness to pay for bundled smart home energy technology and information services, in the context of the Smart Energy Community (ESC) initiative in New York State. ESC is a pilot project launched by the regional electricity and gas provider in Tompkis County, NY. For the estimation of willingness to pay for bundle features, a set of two discrete choice experiments is used to fit logit-type models of demand (McFadden, 1973). Both experiments were part of a survey of residents of upstate New York that took place in 2016, before the installation of 12,400 electric smart meters in 2017. The first discrete choice experiment presented bundled smart energy technology that would help controlling energy use, such as a battery back-up system, a smart thermostat, and a home energy management system. The second discrete choice experiment focused on information provision, including granularity, timing, baseline comparisons, and access method. Both experiments had a price attribute for the additional monthly cost of the added features, which make estimation of marginal willingness to pay possible.

In a surprisingly limited existing literature looking into customer response to smart home energy,[1] this work is closest to Richter and Pollitt (2018) even though the focus of the discrete choice experiments differ. In that paper, using stated-choice survey data collected in 2015 in the UK, the authors fitted a parametric generalized multinomial logit (Fiebig et al., 2010) to model heterogeneity in the willingness to pay for smart electricity

[1] Richter and Pollitt (2018) reviews related demand-side work, including Kaufmann et al. (2013), Dütschke and Paetz (2013), and Paetz et al. (2012). Among the reviewed papers, only (Kaufmann et al., 2013) and (Dütschke and Paetz, 2013) used discrete choice experiments for smart home energy products.

contract terms. Experimental contract attributes were: monthly fee, bill savings, usage monitoring, control of electrical devices, technical support, and data privacy & security. Although British customers are shown to see the value of access to technical support, a statistically significant economic compensation (negative willingness to pay) is found for accepting contract terms involving giving up control, being remotely monitored by the energy provider and sharing usage data with third parties. The unconditional means of the willingness to pay for real-time in-house monitoring with alerts in case of unusual usage and for smart control by the household were not significant.

From a technical point of view, heterogeneity in the willingness to pay for bundled energy products and services is analyzed in this paper using a logit-mixed logit model (Train, 2016), which is an extremely flexible discrete choice model. Although the use of continuous, parametric distributions (as in mixed logit models, Boyd and Mellman, 1980, McFadden and Train, 2000) dominate empirical work including Richter and Pollitt (2018), the use of flexible (semi, non, or seminonparametric) heterogeneity distributions that do not impose a specific shape to the preference variations is desired. The logit-mixed logit model effectively allows the data to reveal the shape of the heterogeneous distribution of willingness to pay measures. Besides, working with parametric mixing distributions is associated with multiple empirical problems (Louviere and Eagle, 2006, Fosgerau and Hess, 2007, Louviere and Meyer, 2008). The logit mixed logit both approximates and generalizes previous discrete choice models with seminonparametric and nonparametric mixing distributions (Bajari et al., 2007, Fosgerau and Bierlaire, 2007, Train, 2008, Bastin et al., 2010, Fox et al., 2011, Fosgerau and Mabit, 2013), many of which exploit polynomial approximations.

The paper is organized as follows. Section 2 proposes a general logit-mixed logit specification that accounts for both observed and unobserved preference heterogeneity (the model as derived in Train, 2016, only considered random preference heterogeneity), partly specified in willingness to pay space. Section 3 describes the data as well as the context and are of study for the empirical application of the proposed flexible logit model. Section 4 discusses estimates of the model, with a focus on willingness to pay for smart home package features and for attributes of a home energy monitoring system. An analysis of sociodemographics that characterize the customer segments that are either more or less likely to opt-in for the smart energy bundles is also presented. Finally, section 5 concludes.

2. A general logit-mixed logit model of preference heterogeneity

2.1. Logit-mixed logit specification

For the estimation of flexible distributions of willingness to pay a logit-mixed logit model (Train, 2016) is derived, taking into consideration both observed and unobserved preference heterogeneity in a utility function that is partly specified in willingness-to-pay space (Train and Weeks, 2005). Let N be the number of customers making discrete choices in the sample. Customer i faces a choice among J alternatives, in each of T time periods. The following general logit-type specification will be considered, in which the customer's truncated indirect utility from alternative j in period t is:

$$
\begin{aligned}
u_{ijt} &= -\sigma_i(\mathbf{x}'_{ijt}\boldsymbol{\omega}_i - p_{ijt}) + \mathbf{d}'_{ijt}\boldsymbol{\delta} + \varepsilon_{ijt} \qquad (1)\\
\boldsymbol{\omega}_i &= \boldsymbol{\Pi}\mathbf{w}_i + \boldsymbol{\epsilon}_i, \qquad (2)
\end{aligned}
$$

where $\mathbf{x}_{ijt}$ and $\mathbf{d}_{ijt}$ are choice-specific attributes, σ_i is the random marginal utility of income of customer i, $\boldsymbol{\omega}_i$ is a random vector of customer-specific willingness to pay for marginal improvements in $\mathbf{x}_{ijt}$, p_{ijt} is price, $\boldsymbol{\delta}$ is a fixed (nonrandom) vector of marginal utilities for characteristics $\mathbf{d}_{ijt}$, ε_{ijt} is an iid type-I extreme value preference shock, $\boldsymbol{\epsilon}_i$ is a random vector of average marginal willingness to pay, $\mathbf{w}_i$ are customer-specific characteristics, and $\boldsymbol{\Pi}$ is a parameter matrix representing observed preference heterogeneity (deterministic taste variations). Note that the system of equations above can be rewritten as a reduced form that involves a combination of random parameters for recovering unobserved preference heterogeneity ($\boldsymbol{\beta}^R =< \sigma_i, \boldsymbol{\epsilon}_i >$) and fixed parameters ($\boldsymbol{\beta}^F =< \boldsymbol{\delta}, \boldsymbol{\Pi} >$):

$$
u_{ijt} = -\sigma_i(\mathbf{x}'_{ijt}\boldsymbol{\epsilon}_i - p_{ijt}) - \sigma_i\mathbf{x}'_{ijt}\boldsymbol{\Pi}\mathbf{w}_i + \mathbf{d}'_{ijt}\boldsymbol{\delta} + \varepsilon_{ijt}. \qquad (3)
$$

Both σ_i and $\boldsymbol{\epsilon}_i$ are assumed to have a discrete heterogeneity distribution leading to the flexible logit-mixed logit specification (Train, 2016) as modified in Bansal et al. (2018) for the consideration of observed preference heterogeneity.[2] Following Train (2016), the discrete heterogeneity distribution of $\boldsymbol{\beta}^R$ is defined using a logit link for the probability w that $\boldsymbol{\beta}_i^R$ equals a specific value $\boldsymbol{\beta}_r^R$ over a support set S, i.e.:

$$
w(\boldsymbol{\beta}_r^R|\boldsymbol{\alpha}) = \Pr(\boldsymbol{\beta}_i^R = \boldsymbol{\beta}_r^R) = \frac{\exp(\mathbf{z}(\boldsymbol{\beta}_r^R)'\boldsymbol{\alpha})}{\sum_{s\in S}\exp(\mathbf{z}(\boldsymbol{\beta}_s^R)'\boldsymbol{\alpha})}, \qquad (4)
$$

where $\boldsymbol{\alpha}$ is a vector of parameters, and $\mathbf{z}(\boldsymbol{\beta}_r^R)$ is a vector-valued function that captures the shape of the mixing distribution and can be specified using the method of sieves

[2] The logit-mixed logit model as originally proposed in Train (2016) considered all parameters to be random.

(i.e. polynomials, step functions or splines). Intuition behind the logit-mixed logit model is straightforward: its discrete preference heterogeneity specification mimics that of a latent class logit model (which is a mixed logit model with a discrete mixture, Boxall and Adamowicz, 2002) with two major differences. First, in a logit-mixed logit model elements $\boldsymbol{\beta}_r^R$ are fixed within a prespecified multidimensional, large-dimensional grid (in a given support set) and are not estimated (whereas the vector $\boldsymbol{\beta}_r^R$ in a latent class logit would be treated as a parameter to be estimated, with a much lower dimensionality). Second, the logit link probability of Eq. 4 is not a class assignment probability but a semi-nonparametric representation of the discrete probability mass at the given point in the grid, which is estimated using the method of sieves.

2.2. Logit-mixed logit choice probabilities

Despite the differences of the logit-mixed logit model with standard mixed logit models, the derivation of the logit-mixed logit choice probabilities still takes advantage of the conditional logit kernel that results from the i.i.d. $EV1$ assumption for ε_{ijt}. If j_{it} denotes the actual choice by customer i at time t, the probability of the sequence of choices $\{j_{i1}, \ldots, j_{iT}\}$ conditional on a realization of the random $\boldsymbol{\beta}_i^R =< \sigma_i, \boldsymbol{\epsilon}_i >$ is:[3]

$$\ell_{i|\sigma_i,\boldsymbol{\epsilon}_i} = \prod_{t=1}^{T} \frac{\exp[-\sigma_i(\mathbf{x}'_{ij_{it}t}\boldsymbol{\epsilon}_i - p_{ij_{it}t}) - \sigma_i\mathbf{x}'_{ij_{it}t}\boldsymbol{\Pi}\mathbf{w}_i + \mathbf{d}'_{ij_{it}t}\boldsymbol{\delta}]}{\sum_{j\in J}\exp[-\sigma_i(\mathbf{x}'_{ijt}\boldsymbol{\epsilon}_i - p_{ijt}) - \sigma_i\mathbf{x}'_{ijt}\boldsymbol{\Pi}\mathbf{w}_i + \mathbf{d}'_{ijt}\boldsymbol{\delta}]}. \tag{5}$$

As in any mixed logit model, the unconditional probability of the sequence of choices made by a customer as a function of the unknown parameters keeps the conditional logit kernel and becomes the individual contribution to the likelihood. In the specific case of the logit-mixed logit model, the unconditional probability of the sequence of choices is simply the following expected value:

$$\ell_i = \sum_{r\in S}\prod_{t=1}^{T} \frac{\exp[-\sigma_{i,r}(\mathbf{x}'_{ij_{it}t}\boldsymbol{\epsilon}_{i,r} - p_{ij_{it}t}) - \sigma_{i,r}\mathbf{x}'_{ij_{it}t}\boldsymbol{\Pi}\mathbf{w}_i + \mathbf{d}'_{ij_{it}t}\boldsymbol{\delta}]}{\sum_{j\in J}\exp[-\sigma_{i,r}(\mathbf{x}'_{ijt}\boldsymbol{\epsilon}_{i,r} - p_{ijt}) - \sigma_{i,r}\mathbf{x}'_{ijt}\boldsymbol{\Pi}\mathbf{w}_i + \mathbf{d}'_{ijt}\boldsymbol{\delta}]} w(\boldsymbol{\beta}_r^R|\boldsymbol{\alpha}). \tag{6}$$

An interesting fact of the expression above is that the weighted average considers all possible values of the random parameters over the support set S (the prespecified multidimensional grid). Thus, the parameters to estimate are reduced to $\boldsymbol{\theta} =< \boldsymbol{\alpha}, \boldsymbol{\Pi}, \boldsymbol{\delta} >$.

[3]Because the logit-mixed logit model a multidimensional grid is prespecified for the random parameters, a realization of the random parameters is equivalent to a random draw from that grid. A specific element $\boldsymbol{\beta}_r^R$ is selected with probability $w(\boldsymbol{\beta}_r^R|\boldsymbol{\alpha})$.

2.3. Maximum likelihood estimator

Adopting a frequentist approach to the estimation of the parameters of interest, the maximum likelihood estimator of the logit-mixed logit model can be derived. The loglikelihood function is constructed from the individual contribution to the likelihood $\ell_i(\boldsymbol{\alpha}, \boldsymbol{\Pi}, \boldsymbol{\delta})$ in Eq. 6:

$$\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\Pi}, \boldsymbol{\delta}) = \sum_{i=1}^{N} \ln \ell_i(\boldsymbol{\alpha}, \boldsymbol{\Pi}, \boldsymbol{\delta}), \tag{7}$$

which is equivalent to

$$\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\Pi}, \boldsymbol{\delta}) = \sum_{i=1}^{N} \ln \left(\sum_{r \in S} \ell_{i|\boldsymbol{\beta}_r^R}(\boldsymbol{\Pi}, \boldsymbol{\delta}) \frac{\exp(\mathbf{z}(\boldsymbol{\beta}_r^R)'\boldsymbol{\alpha})}{\sum_{s \in S} \exp(\mathbf{z}(\boldsymbol{\beta}_s^R)'\boldsymbol{\alpha})} \right), \tag{8}$$

where the product of conditional logit kernels is a function of the fixed parameters and is evaluated at every value in the grid S for the random parameters. Although the loglikelihood does not involve an integral, because the support S is of large dimensions, evaluation and maximization of the loglikelihood is computationally expensive. If there are R random parameters, and for each random parameter a grid of equally-space 1,000 points is considered, then the cardinality of S is 10^{3R}, which becomes explosive quickly. For instance, the case study in Train (2016) has 8 random parameters that result in 10^{24} points in the multidimensional grid.

A solution to the prohibitive computing cost of evaluating Eq. 8 due to the large dimension of S is to work with the maximum simulated likelihood estimator, just as in the standard mixed logit model. A simulated likelihood can be built by considering a random individual-specific subset $S_i \subset S$:

$$\tilde{\mathcal{L}}(\boldsymbol{\alpha}, \boldsymbol{\Pi}, \boldsymbol{\delta}) = \sum_{i=1}^{N} \ln \left(\sum_{r \in S_i} \ell_{i|\boldsymbol{\beta}_r^R}(\boldsymbol{\Pi}, \boldsymbol{\delta}) \frac{\exp(\mathbf{z}(\boldsymbol{\beta}_r^R)'\boldsymbol{\alpha})}{\sum_{s \in S_i} \exp(\mathbf{z}(\boldsymbol{\beta}_s^R)'\boldsymbol{\alpha})} \right). \tag{9}$$

3. Data

3.1. Context and area of study

This study uses data from a survey of homeowners in upstate New York, which was contracted in 2016 by the regional electricity provider to a market research company. The focus of the survey was to collect data on residential customer interest in and response to smart electricity technologies and information services before the Energy Smart Community (ESC) pilot program was launched in Tompkins County, NY. Tompkins County is now the first Energy Smart Community in New York, with a declared goal to study the potential of smart meters and other grid upgrades in increasing energy efficiency and sustainability. In fact, the ESC project is a response to the comprehensive energy strategy for New York Reforming the Energy Vision (REV), which mandates 50% of New York's energy be generated by renewable sources by 2030, as well as to the Energy Roadmap for Tompkins County, which aims at an 80% greenhouse gas reduction from 2008 levels by 2050.

The ESC project in Tompkins County began with the installation of 12,400 electric smart meters and deployment of an advanced grid management system in 2017. In addition to the roll-out of smart meters, the regional electricity provider has also implemented: an online portal (Energy Manager) that allows customers to access their personal (day-by-day, hour-by-hour) energy data and displays customized recommendations to save energy, an online marketplace (Smart Solutions) for energy efficiency products and services, and a price incentive (Smart Usage Plan) to encourage customers to shift their electricity use to off-peak times.

With a population of 101,564 (2010 US Census), Tompkins County comprises the college-town of Ithaca and is home to Cornell University. Cornell researchers have been actually involved in the ESC pilot program, advising the regional utility in terms of community incentives and pricing mechanisms (Khezeli and Bitar, 2017), and collecting data (Bugden and Stedman, 2019) with the goal of "leveraging virtual storage to turn advanced metering infrastructure into a smart service system" (Cornell Chronicle, 2016).

3.2. Data Description

The survey was administered in November 2016. Participants were recruited from the regional utility's email list of residential customers in Tompkins County as well as from a purchased, representative panel sample of residents from the remaining upstate New York areas of the utility's customer base. The final sample comprises 1,093 individuals, with 593 representing the area of interest of Tompkins County.

Table 1 summarizes sociodemographics of the representative sample.

Respondent characteristics	Tompkins County (N=593)	Outside Tompkins (N=500)
Male	48%	48%
18-24 years	13%	11%
25-34 years	14%	15%
35-44 years	16%	16%
45-54 years	20%	20%
55-64 years	17%	17%
65+ years	20%	20%
High school diploma or less	14%	9%
Some college experience	33%	23%
Bachelor's degree	31%	31%
Graduate or professional degree	22%	38%
Household income $< \$25,000$	8%	16%
Household income $\geq \$25,000$ and $< \$35,000$	11%	8%
Household income $\geq \$35,000$ and $< \$50,000$	14%	14%
Household income $\geq \$50,000$ and $< \$75,000$	24%	19%
Household income $\geq \$75,000$ and $< \$100,000$	18%	14%
Household income $\geq \$100,000$	26%	29%
Homeowner	69%	75%

Table 1: Sample Demographic Statistics

3.3. Discrete choice experiments

The survey contained a set of two discrete choice experiments. Each experiment presented 6 choice situations, with 3 unlabelled, alternative bundles and the option to select none at each choice situation. Both experiments included an incremental monthly cost attribute that made possible estimation of marginal willingness to pay for bundle features.

The first discrete choice experiment presented bundled smart energy technology that would help controlling energy use, such as a battery back-up system, a smart thermostat, and a home energy management system. The complete set of bundle features is presented in Table 2.

Bundle features	Levels	Bundle features	Levels
Pricing	$17/month	Length of contract	1-year contract
	$35/month		2-year contract
	$50/month		3-year contract
	$99/month	Home battery storage	No battery back-up system
	$161/month		Battery back-up system
Provider	Regional utility	Internet package	No Internet service
	Local tel/Inet/Cable provider		High-speed Internet (up to 50Mbps)
	Google		High-speeed Inet with streaming
	SolarCity	Home energy management	No energy management system
	Amazon		Smart thermostat
			Connected management system

Table 2: Smart energy products DCE, bundle features and levels

The battery was described as a "system that charges and stores electricity at night for use during the day when electricity is more expensive". As home energy management, the second level was described as "a smart thermostat with a mobile app for controlling your settings", whereas the description for the third level was "a connected home energy management system with a smart thermostat, smart plugs for lighting and appliances, and a mobile app for controlling them all". Table 3 displays a sample of a choice card for this first discrete choice experiment.

If these were your only options, which smart home package would you choose?				
	Bundle A	Bundle B	Bundle C	
Provider	Regional Utility	Google	Your local phone, Internet, cable provider	
Pricing	$50/month	$17/month	$161/month	
Length of contract	1-year contract	2-year contract	3-year contract	
Home battery electricty storage	No battery backup system	No battery backup system	A battery backup system that charges and stores electricty at night for use during the day when electricity is more expensive	
Internet package	No Internet service	High-speed Internet (up to 50Mbps)	High-speed Internet (up to 50 Mbps) with online streaming content membership	
Home energy management	No home energy management system	A smart thermostat with a mobile app for controlling your settings	A connected home energy management system with a smart thermostat, smart plugs for lighting and appliances, and a mobile app for controlling them all	
Preferred choice	○ Bundle A	○ Bundle B	○ Bundle C	○ None

Table 3: Smart energy products DCE, sample choice card

The second discrete choice experiment focused on information provision, including granularity, timing, baseline comparisons, and access method (Table 4).

Information bundle features	Levels
Pricing	Free with pop-up banner ads $1/month $3/month $5/month
Information provided	Usage comparison to same time last year Bill forecasting based on month-to-date and historical usage Usage comparison to similar homes
Electricity usage detail provided	Total electricity usage Total electricity with HVAC usage detail broken out Total electricity with HVAC, water heater, large appliance detail broken out Total electricity detail for HVAC, large appliances, lights and smaller electronics
Usage information timing	Updated once per month Updated daily Real-time
Information access method	Print Online In-home display Phone app

Table 4: Information services DCE, bundle features and levels

An example of a choice card for the second discrete choice experiment is shown in Table 5.

If these were your only options, which home energy monitoring system would you choose?				
	Bundle A	Bundle B	Bundle C	
Information provided	Bill forecasting based on month-to-date and historical usage	Usage comparison to similar homes	Usage comparison to same time last year	
Electricity usage detail provided	Total electricity usage broken out for HVAC and large appliances, as well as lights and smaller electronics like a microwave, hair dryer, etc.	Total electricity with HVAC usage detail broken out	Total electricty usage	
Electricity and/or natural gas usage information timing	Updated daily	Updated once per month	Real time (You turn on something & see the impact immediately)	
Information access method	Phone app	Online	Print	
Pricing	$5/month	$3/month	Free with pop-up banner ads	
Preferred choice	○ Bundle A	○ Bundle B	○ Bundle C	○ None

Table 5: Information services DCE, sample choice card

4. Modeling willingness to pay for smart home energy products and services

4.1. Model specification

Each of the two conjoint questions under analysis considered three alternatives with the possibility of opting out. For each discrete choice experiment the following system of indirect utility – a particular case of the reduced form in Eq. 3 – is specified:

$$u_{ijt} = -\sigma_i(\mathbf{x}'_{ijt}\boldsymbol{\epsilon}_i - p_{ijt}) - \sigma_i\mathbf{x}'_{ijt}\boldsymbol{\Pi}\mathbf{w}_i + \varepsilon_{ijt} \tag{10}$$

$$u_{ojt} = \mathbf{d}'_i\boldsymbol{\delta} + \varepsilon_{ojt}, \tag{11}$$

where Eq. 10 applies to the three bundles of home energy products and services and considers unobserved preference heterogeneity in the willingness to pay for bundle features (recovered in $\boldsymbol{\epsilon}_i$, which is seminonparametrically distributed) as well as observed heterogeneity in the average willingness to pay (taste variations with respect to the population average recovered in the elements of the matrix $\boldsymbol{\Pi}$) as a function of customer covariates ($\mathbf{w}_i$).

For the opt-out alternative, a series of fixed effects is considered in Eq. 11 through a constant and sociodemographics that enter in the vector $\mathbf{d}_i$. As a result, point estimates $\hat{\boldsymbol{\delta}}$ can be used to analyze odd ratios of opting out.

As seminonparametric specification of the distribution of unobserved preference heterogeneity a fourth order polynomial was adopted.

4.2. Willingness to pay estimates

As a result of the adopted model specification, maximum willingness to pay for bundle features are treated as random parameters with a discrete, seminonparametric heterogeneity distribution. From the resulting distribution, point estimates of the average maximum willingness to pay are derived and reported. The model also considers the possibility of customer covariates affecting the average willingness to pay.

Average willingness to pay estimates for the first discrete choice experiment are reported in Table 6, together with the customer covariate effects. In particular, interactions with income, residency in Tompkins County, and generation were able to be estimated. Note that influence of income on willingness to pay is as expected: higher household income is related to a higher average valuation. Residents of Tompkins County are willing to pay more for the battery backup system. Finally, for the other smart energy services there is an important generational divide, with millennials much more likely to see value in the proposed products.

Bundle feature | Segment	Household Income			
	<\$50K	\$50-75K	\$75-100K	>\$100K
Backup battery | Tompkins County	\$40.4	\$42.5	\$48.9	\$50.5
Backup battery | Outside Tompkins County	\$31.3	\$33.0	\$37.9	\$39.1
Smart thermostat | Millennial	\$40.7	\$42.8	\$49.2	\$50.8
Smart thermostat | Generation X	\$12.1	\$12.7	\$14.6	\$15.1
Smart thermostat | Baby Boomer	\$9.5	\$10.0	\$11.5	\$11.9
Energy management system | Millennial	\$47.3	\$49.8	\$57.2	\$59.0
Energy management system | Generation X	\$23.4	\$24.6	\$28.3	\$29.2
Energy management system | Baby Boomer	\$6.7	\$7.0	\$8.1	\$8.4
Highspeed Internet	\$64.8	\$68.2	\$78.4	\$80.0
Highspeed Internet with streaming	\$70.0	\$73.7	\$84.7	\$87.4

Table 6: Smart energy products, average WTP [$/month]

Are these estimates plausible? Whereas the smart energy technologies were not offered at the time of the survey, it is at least possible to contrast the estimates of the resulting willingness to pay for high speed Internet with actual prices of that service. In Tompkins county, it is possible to contract high speed Internet at home for a monthly cost of $40-$80, depending on offered speeds. Estimates of the population average willingness to pay for high speed Internet are in the range of \$65-\$81 per month. The fact that estimates of the willingness to pay for a service that is familiar to customers match actual costs in reassuring as it partly validates customer valuation of the bundle attributes; however, the estimates for the smart energy technologies need to be analyzed with caution as they result from a highly hypothetical scenario at the time of data collection. In fact, when data was collected, information campaigns about the Energy Smart Community pilot program had not yet started.

Besides average willingness to pay, random parameter logit models also provide measures of the variability in preferences. Table 7 summarizes the percentages of the population that according to the seminonparametric estimates exhibit an unconditional positive willingness to pay for the bundle features, paired with the significant interactions with customer covariates.

Bundle feature	Percentage WTP>0
Length of contract	26%
Backup battery | Tompkins County	78%
Backup battery | Ouside Tompkins County	72%
Provider: Regional utility	59%
Provider: Google	30%
Provider: Solar City	31%
Provider: Amazon	31%
Highspeed Internet	100%
Highspeed Internet with streaming	99%
Smart thermostat | Millennial	90%
Smart thermostat | Generation X	65%
Smart thermostat | Baby Boomer	62%
Smart thermostat | Older generations	40%
Energy management system | Millennial	86%
Energy management system | Generation X	70%
Energy management system | Baby Boomer	56%
Energy management system | Older generations	37%

Table 7: Smart energy products, proportion of population with positive preferences

Combining the results of average willingness to pay and its dispersion it is possible to conclude that the most valued attribute in the case of smart home packages was the provision of highspeed Internet. Not only is the willingness to pay for Internet the highest among bundle features, but also there is no negative perception as the whole population is willing to pay a positive amount. For smart energy technology, again the generational divide is patent. On the one hand, both smart thermostats and a home energy management system are highly valued by millennials, with very little negative perceptions. On the other hand, older (than baby boomer) generations fail to see the value of these smart energy options. The backup battery is overall positively valued, with residents of Tompkins County having a slightly higher percentage of customers willing to pay a positive amount for energy storage. Whereas most consumers trust the regional utility, other potential providers are mostly and on average negatively perceived for the provision of smart home products. Longer contracts are disliked.

Regarding willingness to pay for features of a home energy monitoring system, Table 8 presents population averages. Unlike the case of the model for smart home packages, in this second discrete choice experiment it was not possible to identify statistically significant, meaningful interactions with customer covariates.

Bundle feature	WTP	% WTP > 0
Info: bill forecasting based on historical use	-$0.23	44%
Info: usage comparison to similar homes	-$0.78	32%
Detail: HVAC use broken out	$0.39	62%
Detail: HVAC & appliances broken out	$0.78	68%
Detail: HVAC, appliances & small electronics	$0.91	69%
Timing: updated daily	$0.21	57%
Timing: real time	$0.40	59%
Access method: online	$0.83	69%
Access method: in-home display	$0.38	59%
Access method: phone app	$0.04	51%

Table 8: Information services, average WTP [$/month] and proportion of population with positive preferences

Total electricity usage broken out for HVAC, large appliances, lights, and smaller electronics is the most highly valued feature, which is also positively perceived by most customers. Frequent updates (daily or real time vs once per month) are positively valued on average, but the proportion of the population willing to pay a positive amount for these updates is below 60%. Bill forecasting and usage comparison to similar homes are both less preferred than usage comparison to same time last year. In fact, on average customers desire a small compensation for receiving that information. Finally, the preferred access method is online, with an in-home display less favorably perceived (but still positively valued on average). Customers were neutral regarding the use of a phone app, showing indifference with printed information (base level).

4.3. Who is opting in and out?

For the outside option of both discrete choice experiments, fixed parameters $\boldsymbol{\delta}$ in Equation 11 can be used to make inference on odds ratios for opting in or out as a function of customer covariates $\mathbf{d}$. From the point estimates $\exp(\hat{\boldsymbol{\delta}})$, Tables 9 and 10 present the variation of the odds ratio of opting in or out of the energy bundles, respectively for each discrete choice experiment.

In both discrete choice experiments, the odds of choosing none of the bundles are greater for males. In the case of the smart energy products, the odds ratio of opting out are 1.6

Sociodemographic segment	Odds ratio in/out variation
Male	1.6× out
Household with children	3.1× in
Asians	2.6× in
Maximum education: BSc	1.6× in
Maximum education: grad or prof studies	1.5× in
Lives in an apartment	1.5× in

Table 9: Smart energy products bundle, variation in the odds ratios of opting in or out

times higher than that of females, whereas the difference is 1.8 times higher for the information services. The rest of the significant covariates show an increase in the odds ratios of not choosing the outside bundle. For instance, educated households with children are more likely to choose one of the offered bundles.

Even though there was no statistically significant effect of generations on the willingness to pay for features of the home energy monitoring system, it is possible to see a generational divide in the likelihood of choosing one of the bundles in the second discrete choice experiment.

Sociodemographic segment	Odds ratio in/out variation
Male	1.8× out
Household with children	1.9× in
African American	1.9× in
Asian	2.2× in
Maximum education: BSc	1.3× in
Maximum education: grad or prof studies	1.6× in
Baby Boomer	1.1× in
Generation X	1.8× in
Millennial	3.3× in

Table 10: Information services bundle, variation in the odds ratios of opting in or out

For example, the odds ratio of opting in is 3.3 times higher than that of older (than baby boomers) generations.

5. Conclusions

In the context of evolving energy markets and as a response to connect informed customers with cleaner and reliable power, the regional energy provider in upstate New York has launched a Smart Energy Community (ESC) pilot program in Tompkins County. In addition to the installation of smart meters, which is only a first step, the program seeks to develop and test those technologies, markets, and choices that will define the energy utilities of the future. The regional utility has recognized both grid upgrades and community engagement as pillars of the program. In this paper, I have used survey data collected by the regional electric and gas provider before launching the Smart Energy Community program to derive flexible estimates of willingness to pay for smart energy technology and services. The survey included a set of two discrete choice experiments with bundled packages of: 1) smart home technologies (energy storage, and energy management systems) and 2) energy monitoring systems.

To model the stated choices, I have proposed a general logit-mixed logit specification that accounts for both observed and unobserved preference heterogeneity, partly specified in willingness to pay space.

It has been recognized that smart meters alone do not necessarily engage customers in adopting more sustainable energy consumption. One of the goals of the ESC program is to analyze the smart energy potential of energy storage (actual backup batteries or virtual residential storage) that would allow customers to store energy when dynamic electricity prices are low and use that energy when prices are high. Although the flexible estimates show that customers do value the option of residential storage, the average willingness to pay for a backup battery of \$40-\$51 per month by residents of Tompkins County (outside of Tompkins county, the average ranges between \$31 and \$39 as a function of household income) is low: the ownership cost of an installed Tesla Powerwall 2 can reach \$10,000. The logit-mixed logit estimates have also revealed high heterogeneity in the valuation of energy storage: 22% of customers in Tompkins County exhibit a negative willingness to pay for the backup battery (the percentage goes up to 28% outside of Tompkins County). Even though it is often recognized that residents of Tompkins County are environmentally more conscious than those of other areas of upstate New York and the estimates do show a slightly higher valuation of smart energy packages, the difference is rather small.

For home energy management there is an important generational divide with millennials being much more likely to perceive the economic value of smart thermostat and of a home energy management system. For example, 90% of millennials are willing to pay a positive amount per month for having a smart thermostat (only 40% of customers of the silent generation have a positive willingness to pay). For a connected energy management system, the percentage of positive perceptions is 86% among millennials and just 37%

for the silent generation.

Surprisingly, the highest valued feature of a smart home package was access to high-speed Internet. The average willingness to pay for an Internet package not only is highest among bundled attributes ($65-$80 per month), but also appears as positively valued by the whole population. It is important to mention that the estimates obtained match the actual cost of high-speed Internet services, and also that it was the experimental feature most familiar to customers. Since Internet is a high-valued feature of smart home packages, it is interesting to discuss perceptions of who may be providing the service. In addition to the regional energy utility, the first discrete choice experiment included Google, Amazon, the local phone/Internet/cable company, and Solar City as potential providers of smart energy technology services. On average, only the regional utility was positively perceived.

This work has shown that interest for bundled smart energy products and services exists and is economically reflected in a positive average maximum willingness to pay for energy storage, smart thermostats, a fully connected home energy management system, and detailed energy usage that is very frequently updated and can be accessed online. However, said interest in somewhat narrow, with willingness to pay estimates that are low for those technologies that require major investments (such as the purchase and installation of bakcup batteries). There are some limitations in the dataset that was used, including data collection before the launching of the ESC program and a generic experimental design that was proposed by the marketing research company hired by the regional utility. Whereas the estimates are interesting in the context of a population without much knowledge about smart grids and energy markets, at the same time that lack of knowledge means that respondents faced discrete choice games with unfamiliar attributes. Future research in the Tompkins County Smart Energy Community area, since rollout of smart meters and other services has started together with information campaigns, will target informed customers to analyze their perceptions and willingness to pay in the changes and choices that are now available to them. Discrete choice and other behavioral economics experiments will be designed specifically for this particular ESC context.

Acknowledgements

This research is based upon work supported by the National Science Foundation Award PFI:BIC-1632124.

References

Patrick Bajari, Jeremy T Fox, and Stephen P Ryan. Linear regression estimation of discrete choice models with nonparametric distributions of random coefficients. *The American Economic Review*, 97(2):459–463, 2007.

Prateek Bansal, Ricardo A Daziano, and Martin Achtnicht. Extending the logit-mixed logit model for a combination of random and fixed parameters. *Journal of Choice Modelling*, 27:88–96, 2018.

Fabian Bastin, Cinzia Cirillo, and Philippe L Toint. Estimating nonparametric random utility models with an application to the value of time in heterogeneous populations. *Transportation Science*, 44 (4):537–549, 2010.

Peter C Boxall and Wiktor L Adamowicz. Understanding heterogeneous preferences in random utility models: A latent class approach. *Environmental and Resource Economics*, 23(4):421–446, 2002. doi: 10.1023/A:1021351721619.

J Hayden Boyd and Robert E Mellman. The effect of fuel economy standards on the us automotive market: an hedonic demand analysis. *Transportation Research Part A: General*, 14(5-6):367–378, 1980.

Dylan Bugden and Richard Stedman. A synthetic view of acceptance and engagement with smart meters in the United States. *Energy Research & Social Science*, 47:137–145, 2019.

Cornell Chronicle. Cornell researchers aim to unleash 'smart meter' potential, 2016. URL `http://news.cornell.edu/stories/2016/09/cornell-researchers-aim-unleash-smart-meter-potential`.

Elisabeth Dütschke and Alexandra-Gwyn Paetz. Dynamic electricty pricing – which programs do consumers prefer? *Energy Policy*, 59:226–234, 2013.

Denzil G Fiebig, Michael P Keane, Jordan Louviere, and Nada Wasi. The generalized multinomial logit model: Accounting for scale and coefficient heterogeneity. *Marketing Science*, 29(3):393–421, 2010. doi: http://dx.doi.org/10.1287/mksc.1090.0508.

Mogens Fosgerau and Michel Bierlaire. A practical test for the choice of mixing distribution in discrete choice models. *Transportation Research Part B: Methodological*, 41(7):784–794, 2007.

Mogens Fosgerau and Stephane Hess. Competing methods for representing random taste heterogeneity in discrete choice models. Technical report, Working paper, Danish Transport Research Institute, Copenhagen, 2007.

Mogens Fosgerau and Stefan L Mabit. Easy and flexible mixture distributions. *Economics Letters*, 120 (2):206–210, 2013.

Jeremy T Fox, Stephen P Ryan, and Patrick Bajari. A simple estimator for the distribution of random coefficients. *Quantitative Economics*, 2(3):381–418, 2011.

Simon Kaufmann, Karoline Künzel, and Moritz Loock. Customer value of smart metering: explorative evidence from a choice-based conjoint study in Switzerland. *Energy Policy*, 53:229–239, 2013.

Kia Khezeli and Eilyan Bitar. Risk-sensitive learning and pricing for demand response. In *IEEE Transactions on Smart Grid*, 2017.

Jordan Louviere and Thomas Eagle. Confound it! that pesky little scale constant messes up our convenient assumptions. In *Proceedings of the Sawtooth Software Conference*, pages 211–228, 2006.

Jordan J Louviere and Robert J Meyer. Formal choice models of informal choices. 2008.

Daniel McFadden. Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, pages 105–142, 1973.

Daniel McFadden and Kenneth Train. Mixed mnl models for discrete response. *Journal of Applied Econometrics*, 15(5):447–470, 2000.

Alexandra-Gwyn Paetz, Elisabeth Dütschke, and Wolf Fichtner. Smart homes as a menas to sustainable energy consumption: a study of consumer perceptions. *Journal of Consumer Policy*, 35(1):23–41, 2012.

Laura-Lucia Richter and Michel G. Pollitt. Which smart electricity service contracts will consumers accept? The demand for compensation in a platform market. *Energy Economics*, 72:436–450, 2018.

Kenneth Train. Mixed logit with a flexible mixing distribution. *Journal of Choice Modelling*, 19:40–53, 2016.

Kenneth Train and Melvyn Weeks. Discrete choice models in preference space and willingness-to-pay space. In Riccardo Scarpa and Anna Alberini, editors, *Applications of Simulation Methods in Environmental and Resource Economics*, volume 6 of *The Economics of Non-Market Goods and Resources*, pages 1–16. Springer Netherlands, 2005.

Kenneth E Train. EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling*, 1(1):40–69, 2008.