WORKING PAPER SERIES
NO. 18006

INFORMATION VERSUS AUTOMATION AND IMPLICATIONS FOR DYNAMIC PRICING

BRYAN K. BOLLINGER
Fuqua School of Business, Duke University

WESLEY R. HARTMANN
Graduate School of Business, Stanford University

OCTOBER 24, 2018

Hoover Institution Working Group on
Intellectual Property, Innovation, and Prosperity
Stanford University

www.hooverip2.org

Information versus Automation and Implications for Dynamic Pricing

Bryan K. Bollinger
Fuqua School of Business
Duke University

Wesley R. Hartmann
Graduate School of Business
Stanford University

September 26, 2018

Abstract

Essential resources like electricity and water can experience rapidly changing demand or supply while the other side of the market is unchanged. Short-run price variation could efficiently allocate resources at these critical times, but only if consumers exhibit short-run demand elasticity. The question for firms in these markets has always been how to enable this response. Randomized control trials are increasingly used to test dynamic pricing and technologies that can assist in response by providing information and/or automated response. But, the trials typically do not randomize short-run prices. This paper illustrates how demand from a randomly assigned control group can be used to test the effectiveness of different technologies in increasing short-term price elasticity. To do so, we use a non-parametric control function approach that eliminates the bias inherent in estimating short-term price response using only household random assignment. We find that only automation technology leads to the short-term price elasticity needed to justify real-time pricing.

Automation Technology, Demand Response; Short-run Elasticities; Utilities; Energy; Electricity; Field experiments.

1 Introduction

The demand and supply of essential resources such as electricity and water regularly fluctuate thereby motivating frequent price changes. However, prices in these markets have traditionally been fixed or, if variable, predetermined. Without a price mechanism in place to align firm and consumer incentives, firms' typical conservation strategies only incentivize long-run behavioral changes, such as providing rebates on efficient appliances. Such policies lead to over-consumption by non-adopters during times of short-run scarcity and under-consumption in times of abundance by adopters.

There has been a push to reconsider dynamic pricing as a mechanism to induce short-run demand response in electricity markets where demand and supply are both highly variable (Joskow and Wolfram, 2012). Potential efficiency gains from dynamic pricing are clear (Borenstein and Holland, 2005), but in the absence of assisting technologies, previous evidence suggests that utility consumers exhibit inelastic demand (Reiss and White, 2005; Allcott, 2011; Ito, 2014). Recent research has shown that communication technology that provides information on real-time prices leads to demand reductions, but such reductions are better characterized as long-run behavioral changes (Jessoe and Rapson, 2014; Faruqui, Sergici, and Sharif, 2010). Thus the extent to which communication technology facilitates short-term price response is unclear.

One challenge in measuring short-run price response is that previous research evaluated field experiments that used (long-term) household random assignment across price distributions (Jessoe and Rapson, 2014; Harding and Lamarche, 2016; Liu, Ni, and Shen, 2017). Work in marketing and economics makes clear, however, that varying price distributions is ideal for estimating long-run elasticities (Mela, Gupta, and Lehmann, 1997; Erdem, Imai, and Keane, 2003; Hendel and Nevo, 2006). Experimental measurement of short-run elasticities requires randomization at the individual X time level. While we are unaware of any such experiments, we illustrate how demand from a randomly assigned control group exposed to a fixed price can allow researchers to non-parametrically condition on the unobserved demand shocks that would otherwise bias short-run elasticities.[1]

With the ability to separately measure long-run and short-run elasticities in an experimental setting, we test whether information technology can increase short-run demand elasticity or whether automation technology is required. Automation technology addresses the problem that even informed consumers find it too costly to regularly alter their electricity consumption. We analyze data from a field experiment run by a large electric utility in the southern United States that used an online portal to provide information, an in-home display (IHD) to provide communications, and a programmable communicating thermostat (PCT) to provide automation. The PCT is a traditional, digital thermostat with the augmented ability to automatically respond to price changes based on the consumer input of how they tradeoff comfort and savings.[2] As speculated in Allcott (2011), we do indeed find that the PCT provides significant improvements in demand response.[3] Although both information and automation are effective at reducing long-run demand, automation technology can uniquely generate the short-run demand elasticities essential to facilitate dynamic pricing in electricity markets.

Our estimates reveal that for all treatments without automation technology, the primary (and in most specifications exclusive), effect of dynamic pricing was a permanent reduction in

[1]To separately identify the short-run demand elasticity, we derive how the randomly assigned control group with a fixed price schedule can be used to form a non-parametric control function that eliminates confounds in estimating short-run demand response for those on the dynamic schedule. The logic is as follows: Both the treatment and control group will be exposed to the same short-run shocks to utility, e.g. they experience the same variation in local temperature. Responses to these shocks may differ, but if each treatment group's response to unobserved utility shocks monotonically transforms into demand shocks, there exists a monotonic function mapping unobserved determinants of control group demand to unobserved determinants of demand by the treated groups. For this control-function approach to be valid, potentially confounding unobservables must relate to price through a scalar index. For example, in the auctions literature a reserve price (Roberts, 2013) or the number of bidders (Haile, Hong, and Shum, 2003; Guerre, Perrigne, and Vuong, 2009; Compiani, Haile, and Sant'Anna, 2017) reflect confounding unobservables. In the case of dynamic electricity pricing, an index of the potential gap between supply and demand should drive the price. In our appendix, we illustrate and assesses the problem that can arise when multiple unobservables (e.g. a high and low temperature) differentially affect price and demand and show even then that there are minimal implications for our elasticity estimates.

[2]When a PCT is installed, consumers are assisted in setting up the device to reflect their preferences.

[3]Cappers, Goldman, and Kathan (2010) show that demand response without home automation has increased by 10% since 2006 in reducing peak load, and that existing demand response resource potential ranges from 3 to 9% of a region's summer peak demand in most regions.

electricity consumption, shown by a leftward shift of a perfectly inelastic demand curve. The non-parametric approach described above finds significant differences in demand between two price points for only 5 of 16 states (combinations of price and levels of the unobservable demand shock), for the IHD, whereas for households with the PCT, significant price effects exist in all but one of the 16 states. Furthermore, the estimated short-run price responses for automation technology are three or more times greater than without. We also compare these results to those from linear models, fixed effect models and a parametric control function model that applies the same identifying intuition of our non-parametric approach but uses a traditional two-step sequential linear regression estimation. We show that the control function is necessary to remove confounds in identifying short-run demand response.

Our results raise serious concerns for implementing dynamic pricing in settings where consumers do not have automated ways of responding to variation in generally low prices (such as those observed in resource markets). The long-run demand responses delivered by these pricing policies may be better generated by communication outreach and/or appliance rebates that are intended for long-run demand reductions, as opposed to exposing consumers to prices they are unable or unwilling to respond to when faced with regular adjustment costs. Short-run response is, however, needed in industries where demand and supply conditions can rapidly change. Our results reveal that automation technology meets Borenstein and Holland (2005)'s benchmark of even a moderate price response being sufficient to increase efficiency.

Though not the focal point of our study, some may question whether automation can sufficiently replicate consumer preferences to develop welfare improvements. Welfare calculations can be tricky, especially in markets such as this where pollution externalities may be present. One encouraging starting point to a welfare improvement is the recognition that a household given a PCT can always set it to mimic a traditional thermostat. Setting aside the costs of installing and inputing preferences into a PCT, it should weakly improve consumer surplus. Ignoring equilibrium price response to the technology, we use the non-parametric demand curves estimated above to estimate changes in consumer surplus under some basic assumptions. We find that adding a PCT to a portal and IHD-equipped household increases welfare by \$27 in the summer of 2011. At a cost of \$250, it would take roughly nine years for a consumer to break even. Adoption of automation may therefore rely on incentives from the utility which realizes benefits from delayed capacity investment of more than \$800 over the lifetime of an automation device.[4]

These findings suggest an important role for automation technology within both the literature Stigler (1961) initiated about limited price response and the energy economics literature. Information economics focuses on buyers incurring search costs to resolve their uncertainty, usually regarding sellers' prices. Search costs may be high[5], but we find that even if a customer has price and consumption displayed in their home, it requires automation technology to actually read and respond to short-term price changes. Our findings also document the potential for automation technology to move utilities to dynamic pricing plans. In addition, such technologies are increasingly important in many domains: they can reduce or eliminate buyers' costs of acquiring, processing, and responding to the abundance of information inundating consumers today.

The paper proceeds as follows. In Section 2 we describe the experimental treatments, and in Section 3 we provide descriptives of our unique data set. In Section 4, we estimate average treatment effects that illustrate the total demand reductions from combinations of technology and flexible pricing assignments. In Section 5, we then examine whether these demand reductions are due to long-term behavioral responses to treatment assignment or whether the demand reductions occur because of responses to short-term price variation. We first provide a discussion of the bias that results from using the price variation across randomly assigned price schedules to identify short-term price response; we further present our

[4] We were informed by consultants for the utility that a permanent 1kw decrease in critical period demand has a net present value of \$700 based on delayed capacity investment. The treatment effects we estimate for PCT devices range from 1.15 to 1.17 kw reductions during critical periods.

[5] Recent research documents large search costs across markets as diverse as online insurance quotes (Honka, 2014) to the aisles or shelf facings in a grocery store (Seiler, 2013).

Figure 1: In-Home Display (IHD)

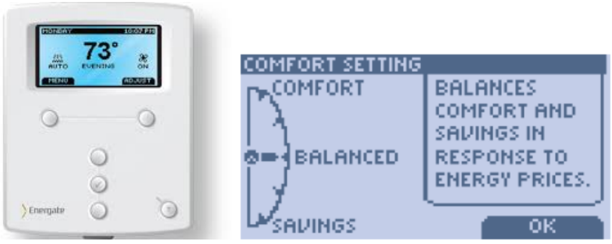
estimation approach to control for demand endogeneity while still leveraging inter-temporal variation in prices. We follow with our demand estimates, and a calculation of the consumer surplus gain from home automation. We provide concluding remarks in Section 6.

2 Research Design

A large utility in the southern United States conducted the experiment we study in the summer of 2011. Consumers in two cities received letters informing them of the program and were told that they could see a reduction in their utility bill for participating. Households were given a best bill guarantee such that if their energy cost exceeded what it would have been from not participating, they would pay that alternative amount. This helped to increase the participation rate while still ensuring incentive compatibility. In Appendix F we confirm that this incentive did not distort our findings. Smartmeter-enabled technologies were provided to all consumers in an attempt to help them monitor and reduce their consumption. Conditional on participating, households were assigned to one of nine different conditions, including a control condition.

Aside from the control condition, the treatment conditions combine one of four technology treatments with one of two pricing plans (designed to keep average bills approximately the same without adjustments to usage). The four technology treatments are: i) a web portal to monitor usage and price, ii) the IHD, which removes the need to be online to assess usage and price, iii) the PCT, which can be set to turn off air conditioning based on the time,

Figure 2: Programmable Communicating Thermostat (PCT)



in-home temperature, or price, and iv) a combination of all three technologies (All3). See Figures 1 and 2 for pictures of the IHD and PCT technologies.

The pricing treatments are i) time of use pricing (TOU) which sets different prices for the peak and off-peak hours on weekdays, and ii) variable peak pricing (VPP), in which the peak price can be varied by the utility depending on its aggregate demand. The price schedules are described below:

- ***Time of Use with Critical Pricing:***
    - Off-peak price of 4.2c per kWh: Weekends, off-peak times on weekdays and holidays.
    - On-Peak price of 23c per kW: On-peak times (2pm-7pm) on weekdays.
    - Critical price of 46c per kWh: Rare price overcall for 2 to 8 hours at any time during the year. Requires 2 hours notice.

- ***Variable Peak Pricing with Critical Pricing:***
    - Off-peak price of 4.5c per kWh: Weekends and off-peak times on weekdays.
    - On-Peak prices, communicated by 5pm the day before (average of 13.6c per kWH):

Price / kWh	4.5	11.3	23	46
Freq in Days	50	37	23	10

    - Critical price of 46c per kWh: Rare price overcall for 2 to 8 hours at any time during the year. Requires 2 hours notice.

- ***Control Pricing:***

    - 8.4c per kWh regardless of day or time.
    - Increases to 9.68c after 1,400 cumulative kWh in the month.

To aid with external validity, we note that Henley and Peirson (1994) also observe TOU pricing that increases prices by an order of magnitude (from between 2.22 and 9.71 off-peak to 40.71 peak). In Herter, McAuliffe, and Rosenfeld (2007), the critical peak price (CPP) price was three times the TOU price, whereas it is double in our scenario, and CPP prices are 50 or 68 cents in Herter and Wayland (2010), similar to our value of 46 cents.

The randomization occurred at the account level rather than the meter level. Every household received their treatment sometime between April 2, 2011 and May 4, 2011, and they remained in their assigned treatment throughout the entire summer (and afterwards). We restrict the analysis to accounts with a single meter to avoid the possibility that demand response is spread across multiple meters and potential multiple households (in the case such accounts may be for multi-family homes with multiple meters). We exclude from the analysis households who have the low-income price rate (all assigned non-randomly to the control group), and those without a treatment start date recorded. Since the vast majority of homes have air conditioning, we exclude those without. If we were to include them, they would have to be analyzed separately because households without air conditions were randomly assigned between the portal and IHD treatments (air conditioning is required for a PCT). Finally we drop the remaining two households without survey data on income and age, and the two without 2010 usage data.

Table 1: Summary Statistics

		N	Mean	Std. Dev	p25	p75
Avg. Usage 2011	Total	2,210	2.400	1.112	1.656	2.979
	Critical	2,170	3.031	1.717	1.810	4.023
	Peak	2,210	2.940	1.479	1.934	3.742
	Off-Peak	2,210	2.305	1.078	1.563	2.847
Family		2,210	0.394	0.489		
Young		2,210	0.333	0.472		
Mature		2,210	0.273	0.446		
Low Income		2,210	0.274	0.446		
Middle Income		2,210	0.290	0.454		
High Income		2,210	0.436	0.496		

3 Data

The final dataset includes 12,535,790 hours of consumption data between June 1 and October 1 in both 2010 and 2011, from 2,210 households. [6] Table 1 reports summary statistics for the households in the final sample. Electricity usage in 2011 is broken down into three categories: critical, peak, and off-peak. Average usage in the critical and peak periods is about 3 kW with a larger standard deviation for the critical period[7], which typically represents a hot summer day where demand threatens to exceed supply. Off-peak consumption is 2.305 kW with a standard deviation of 1.078. 39% of the households are families with the others split between younger and mature households. 44% of households are high income with low and middle income representing 27% and 29% of the sample, respectively.

3.1 Randomization and Compliance

Of the 2,210 households, 173 were not given their original technology treatment assignment. All households received the price schedule to which they were randomly assigned. Table 2 shows the adherence to the treatment assignment - almost all deviations occurred due to

[6] 29 hourly observations were dropped in which the recorded consumption exceeded 30 kW, with values ranging between 123 and 59238 and thus were clearly data recording errors.

[7] There are fewer observations for the critical periods because some households are not observed during the rare critical events.

Table 2: Technology Assignment and Technology Received

		Technology Received					
		Control	Portal	IHD	PCT	All3	
	Control	317	0	0	0	0	317
Technology	Portal	0	523	0	0	1	524
Assigned	IHD	0	28	416	0	3	447
	PCT	0	70	3	429	2	504
	All3	0	29	31	6	352	418
		317	650	450	435	358	2,210

different forms of installation issues which led to being "downgraded" to the portal condition. Differences in the bin sizes are due to missing usage data and those households were discarded by the utility before they supplied the data.

In Table 3a, we demonstrate the covariate balance across households for the assigned treatments. We do the same for the received treatments in Table 3a. There are no significant differences in the 2010 hourly usage across treatments (assigned or received), and we also see comparable balances across the demographic bins for age and income.[8] There are slightly more households in the "high" income bin and fewer in the "medium" bin for the portal condition. This difference, however, is small. In Appendix A, we verify the validity of the randomization by testing whether there are any significant pre-treatment outcome differences between the treatment and control groups and find no significant differences.

As in Harding and Lamarche (2016), we are unable to compare enrolled households from non-enrolled households. However, in terms of generalizability, we do not need these results to extrapolate to non-enrolled households. What is paramount to the utility is the enrollment of households into a program with dynamic pricing and the short-term price response from this particular set of households. These are the households which will allow the utility to match short term fluctuations in demand and supply with a price instrument.

[8]The age variable reports whether the household is a family with kids, and the relative age of the household head otherwise.

Table 3: Covariate Balance

(a) Assigned Treatments

Variable		Control	TOU Portal	TOU IHD	TOU PCT	TOU All3	VPP Portal	VPP IHD	VPP PCT	VPP All3
Prior	Mean	2.208	2.236	2.172	2.150	2.318	2.333	2.163	2.200	2.290
Usage	S.D.	1.253	1.025	1.071	0.921	1.100	1.187	1.102	1.105	1.202
	Low	0.322	0.229	0.278	0.247	0.273	0.273	0.290	0.281	0.263
Income	Med	0.293	0.241	0.323	0.298	0.297	0.233	0.286	0.317	0.335
	High	0.385	0.530	0.399	0.455	0.431	0.495	0.424	0.402	0.402
	Family	0.192	0.257	0.305	0.310	0.278	0.265	0.304	0.285	0.292
Age	Young	0.498	0.394	0.404	0.341	0.388	0.396	0.357	0.386	0.340
	Mature	0.309	0.349	0.291	0.349	0.335	0.338	0.339	0.329	0.368

(b) Received Treatments

Variable		Control	TOU Portal	TOU IHD	TOU PCT	TOU All3	VPP Portal	VPP IHD	VPP PCT	VPP All3
Prior	Mean	2.208	2.173	2.206	2.207	2.309	2.241	2.221	2.291	2.250
Usage	S.D.	1.253	1.014	1.101	0.904	1.092	1.199	1.089	1.105	1.193
	Low	0.322	0.258	0.283	0.218	0.261	0.294	0.295	0.260	0.241
Income	Med	0.293	0.239	0.323	0.324	0.288	0.244	0.263	0.329	0.362
	High	0.385	0.503	0.394	0.458	0.451	0.462	0.442	0.411	0.397
	Family	0.192	0.274	0.319	0.287	0.272	0.271	0.286	0.288	0.310
Age	Young	0.498	0.352	0.407	0.375	0.402	0.374	0.375	0.388	0.345
	Mature	0.309	0.374	0.274	0.338	0.326	0.356	0.339	0.324	0.345

3.2 Usage by Treatment

Before performing any regression analyses to determine treatment effects, we summarize the mean consumption in the critical, peak, and off-peak periods for the different treatments received in Table 4. These averages are performed across household-hour observations for the treatment summer (2011). There are virtually no differences when using treatments assigned instead, as shown in the Appendix in Table 10.

There is a clear difference in critical period consumption across technology treatments and price treatments. Looking at the treatment received, in the TOU price condition, critical consumption drops by 0.391 kW in the portal condition (relative to the control), and only an additional 0.087 kW with the addition of the IHD. The PCT drops consumption by an additional 0.773 kW over the portal alone. All3 actually leads to more consumption than the PCT in isolation, presumably due to more overrides of the automated setting. For the VPP price condition, we see that the portal alone leads to a drop on average of 0.286 kW. The IHD leads to a further reduction over the portal of only 0.051 kW, whereas the PCT leads to a reduction of 0.828 over the portal alone. This evidence shows a large impact of peak consumption reduction for the PCT and virtually none for the IHD, consistent with Harding and Lamarche (2016).

There is a small decline in non-critical peak usage when including the IHD with the portal (although it is only present for the treatment received), and again we see large declines in consumption in both the PCT and All3 treatments. Consumption is lower for all of the TOU technology treatments relative to the same treatment in VPP pricing, although we cannot rule out that this effect is due to the fact that the VPP prices were smaller on average over the course of the entire summer.[9] There also is a slight increase in off-peak consumption for the PCT and All3 treatment conditions in both price treatments, which is consistent with

[9] To test this, we regressed usage on price and treatment using hour-in-data dummy variables to control for any common demand shock, which then uses only the price variation across experimental price treatments, and found that there was no significant effect in the technology treatments across pricing conditions after controlling for price.

Table 4: Hourly Usage by Technology Received

(a) Critical Period Hourly Demand by Technology Received

control	price	portal	IHD	PCT	All3
3.692 (2.276) N=311	TOU	3.301* (2.094) N=307	3.214* (2.022) N=222	2.528*** (1.940) N=213	2.721*** (1.980) N=180
	VPP	3.406 (2.188) N=331	3.355† (1.984) N=216	2.578*** (2.040) N=219	2.540*** (2.000) N=171

(b) Non-Critical Peak Period Hourly Demand by Technology Received

control	price	portal	IHD	PCT	All3
3.339 (2.249) N=317	TOU	3.046† (2.041) N=310	3.003† (1.980) N=226	2.443*** (1.934) N=216	2.617*** (1.982) N=184
	VPP	3.180 (2.187) N=340	3.136 (1.969) N=224	2.735* (2.040) N=219	2.765** (2.018) N=174

(c) Non-Critical Off-Peak Period Hourly Demand by Technology Received

control	price	portal	IHD	PCT	All3
2.303 (1.902) N=317	TOU	2.209 (1.759) N=310	2.270 (1.794) N=226	2.310 (1.761) N=216	2.440 (1.865) N=184
	VPP	2.316 (1.901) N=340	2.323 (1.791) N=224	2.456 (1.900) N=219	2.404 (1.891) N=174

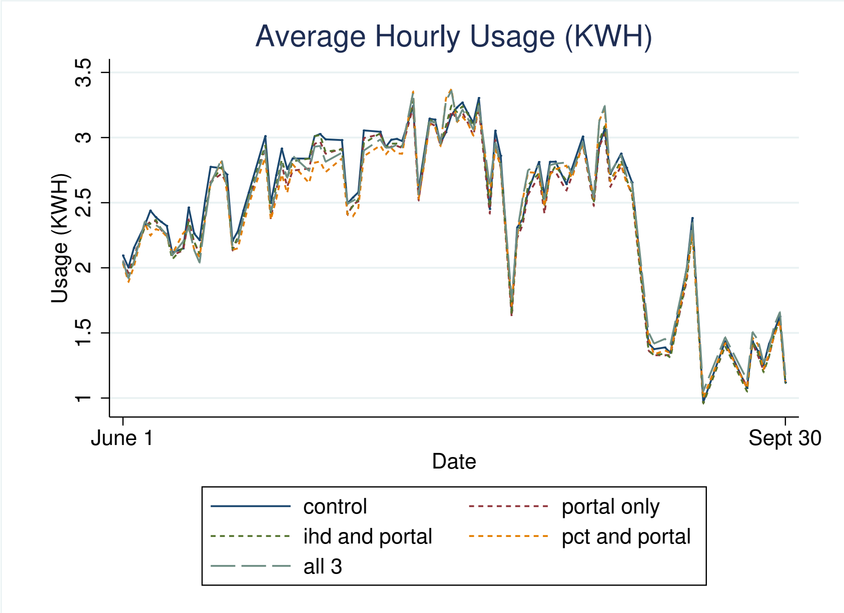
Standard deviation in parentheses. Stars indicate significant differences in means relative to the control group.
*** p=0.001, ** p=.01%, * p=.05. †p=.10

consumption patterns we show later in Figure 5 in which the air conditioning needs to work more in the post-peak period after the PCT turns off the air conditioning in the peak period, automating inter-temporal substitution.

3.3 Descriptive Patterns in the Data

In this subsection, we provide depictions of the data to illustrate three things: i) electricity usage does exhibit seasonal and temporary shocks suggesting value to dynamic pricing, ii) households treated with dynamic pricing do reduce critical and peak consumption relative to control households with greater reductions occurring for PCT enabled households, and iii) households with PCTs exhibit clear short-run price response whereas IHD and portal

Figure 3: Average Hourly Electricity Consumption by Treatment



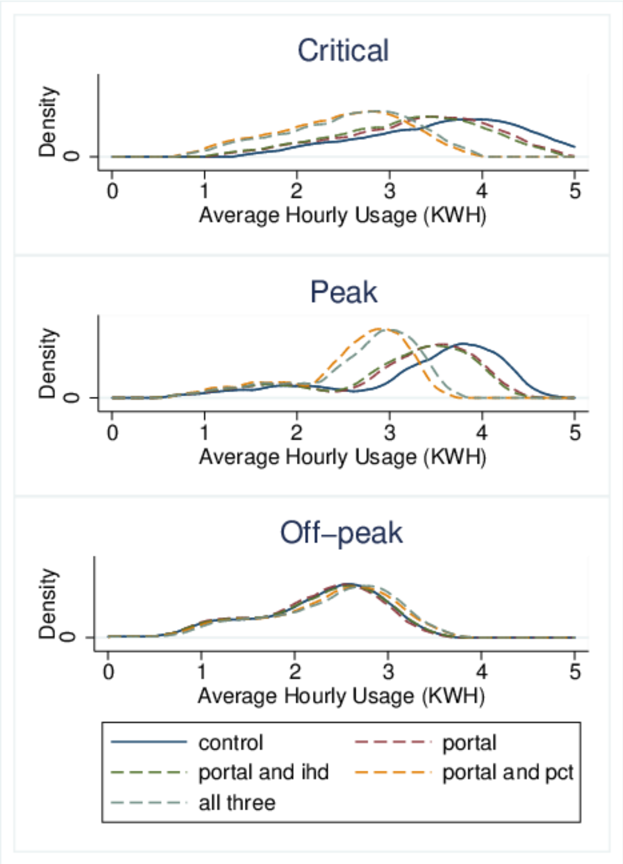
enabled consumers have more gradual responses to price changes.

Figure 3 shows the household average hourly usage by treatment across the days of summer 2011. The seasonality of electricity use is clear as the usage increases through the middle of summer and declines while approaching the end of September. However, the pattern is not smooth, exhibiting substantial downward and upward spikes in electricity demand. These are exactly the types of spikes that motivate the use of dynamic pricing to reallocate demand when scarcity suddenly rises or falls. It is also clear that all the technology treatments (with TOU or VPP pricing), do lead to an aggregate reduction in usage since the control group demand is visibly above that of the others.

To look further into the differences in usage by treatment, Figure 4 shows kernel density estimates of the average hourly usage, in which the averages are taken by household over the critical events hours, and non-critical peak and off-peak hours of each day. There is a drop for all technology treatments during the critical and peak hours, but a small increase in usage during the off-peak hours. The PCT and All3 technologies reduce peak demand far more than the portal and IHD.

In addition to generally greater reductions in demand for the PCT devices, these con-

Figure 4: Average Hourly Electricity Consumption by Treatment, on Peak Days

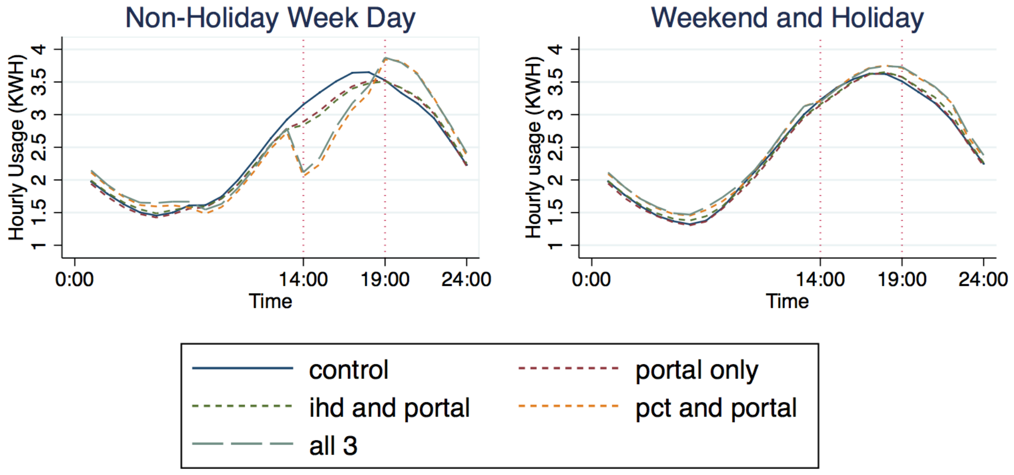


sumers are also clearly more responsive to price movements. Notably, the peak vs off-peak pricing generates a discontinuity in price at 2pm (14:00 hours) on every weekday. Short-run price response should generate a corresponding discontinuity in consumption at the same time. Figure 5, plots average hourly usage by treatment for days with peak periods (non-holiday weekdays), versus those without (weekends and holidays when no discontinuity in price exists). The superiority of the PCT device in responding to short-run price changes is evident with the sudden reduction in electricity consumption at the beginning of the peak period. The IHD and portal-assigned households exhibit only gradual divergence from the control group demand beginning at 2pm.[10] While the discontinuous price change allows for a useful evaluation of price response, it should be noted that a wider rollout of PCT devices would call for a smoothing of the start times to avoid the sudden demand shock this would otherwise require the utility to accommodate with additional electricity generation.

To test the significance of these differences, and to control for any other factors not addressed by the randomization and which might affect our estimates , we use regression analysis in the next section to estimate the treatment effects for demand reductions in critical and peak time periods. We focus on the treatment received since there appear to be few differences from treatment assigned.

[10]Interestingly, IHD and portal households with time-contingent thermostats could set a different preferred temperature at 2pm when the known discontinuous price increase occurs. This suggests there might be alternative treatments that could be targeted at educating or enabling customers to set their thermostats accordingly. We do not know the costs of such an intervention and have no effect estimates for comparison but a clear advantage of the PCT is that if the utility ever changed the timing of price movements, an additional costly intervention would not be required.

Figure 5: Average Hourly Electricity Consumption by Treatment



4 Average Treatment Effects

4.1 Regression Analyses

To estimate the treatment effects for our panel of electricity consumption in the summer following treatment, we use the following simple regression:

$$y_i = \alpha_0 + \alpha_1 \tilde{A}_i + \epsilon_i, \tag{1}$$

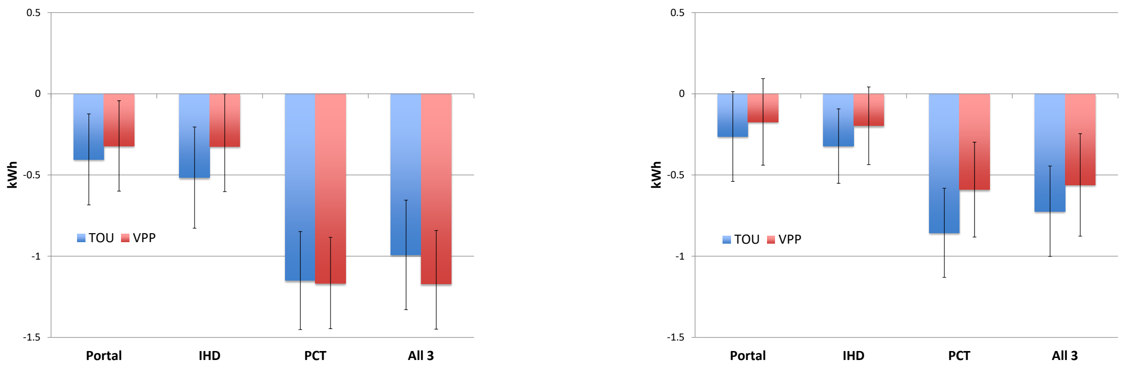
where y_i is household i's average hourly electricity consumption in either critical, peak, or off-peak periods – we run separate regressions for each. We collapse all of the data across time within these three periods and use bootstrap standard errors (across households) to calculate the most conservative standard errors. $\tilde{A}_i$ is a vector of dummies corresponding to each of eight treatments (the four technology treatments interacted with the two price treatments), with zeros for all but the household's treatment A_i. α_0 captures the average usage of the control group, while α_1 is a vector which measures the change in usage attributable to each treatment. ϵ_i is the unobservable consumption shock, which is uncorrelated with treatment because of the randomization. The estimates of α_1 are plotted separately for critical and peak periods in Figure 6, and off-peak are plotted in Appendix B. All coefficients and standard errors can also be found in Table 5. As robustness checks, Appendix B contains estimates of regressions with various sets of controls, using the uncollapsed hourly data to identify the treatment effects.

We begin our discussion of the treatment effects for the critical time periods. The critical period demand reductions shown in Figure 6a are statistically significant for all technology X price treatment combinations. Under VPP, both technology treatments reduce demand by 0.32 kWh, while TOU-IHD is a slightly greater reduction at 0.52 relative to 0.40 for the TOU-Portal. The IHD reductions we estimate of 0.516 and 0.321 kW in TOU and VPP pricing, respectively, correspond to decreases of 14% and 8.8%. This is consistent with the reduction from IHD found in Jessoe and Rapson (2014) of 8-22% during critical pricing events. The comparable reductions with the portal and IHD imply that the increased price salience from the IHD (the portal requires the user to log in daily to learn the upcoming price information), is not sufficient to overcome the adjustment costs. It may therefore be possible that the portal technology was sufficient to inform customers during these time periods. Turning to the PCT, we find that reductions are over 1.15 and 1.17 kW for TOU and VPP price treatments respectively. The All3 demand reduction is the same for VPP, but slightly lower at 0.99 kWh for TOU. The results indicate a clear superiority of the PCT technology in delivering long-run demand reductions, in addition to any effect on short term price elasticity.

Figure 6b demonstrates significant demand reductions (as indicated by the 95% confidence bars), for all technology X price treatments in non-critical peak periods as well. The only exception is the portal and IHD with VPP, although the treatment effects are also not significantly different than when paired with TOU pricing. The two treatments that include the PCT device to automate adjustment to temperature and price again exhibit statistically larger reductions in peak demand than the portal and IHD, for both pricing conditions. The point estimates for the reductions under TOU pricing are all lower than under VPP (for the

Figure 6: Demand Reductions by Treatment

(a) Critical (b) Peak



Error bars indicate 95th percent confidence intervals. For both Critical and Peak, Portal and IHD differences are not statistically significant; PCT and All3 are both significantly different from IHD; PCT and All3 differences are not statistically significant.

same technology), but this can be explained due to the fact that the realized prices were on average higher under TOU pricing than under VPP (the average realized peak price for TOU is 23c, whereas it is 13.6c in VPP). Since all future TOU peak prices are known at the time of assignment, it could also be the case that both Portal and IHD households with TOU pricing permanently set their thermostats in response to their assigned plan, and that there is no value added from having an in-home display of price information that was previously communicated to follow an easily recalled pattern.

None of the technology X price treatments exhibit statistically significant changes during off-peak periods. They are shown in Appendix C, Table 10. Customers with home automation (those with the PCTs and All3 technology treatments), exhibit mild increases in usage, which is consistent with Figure 5's display of greater demand immediately following the end of the peak hours.

One potential critique of these results is that we use treatment received when there is a small fraction of households who do not receive their assigned treatment. We have previously presented evidence showing comparable consumption by treatment when using both treat-

Table 5: Treatment Effect Regression Results, Collapsed Data

Treatments		Critical estimate	Critical s.e.	Peak estimate	Peak s.e.	Off-Peak estimate	Off-Peak s.e.
TOU	**Portal**	-0.404**	(0.135)	-0.263	(0.141)	-0.078	(0.086)
	IHD	-0.516**	(0.172)	-0.322**	(0.117)	-0.035	(0.103)
	PCT	-1.150***	(0.154)	-0.856***	(0.140)	0.032	(0.081)
	All3	-0.992***	(0.162)	-0.723***	(0.142)	0.118	(0.096)
VPP	**Portal**	-0.321*	(0.156)	-0.173	(0.136)	-0.007	(0.094)
	IHD	-0.324*	(0.160)	-0.197	(0.122)	0.016	(0.117)
	PCT	-1.168***	(0.140)	-0.590***	(0.149)	0.153	(0.085)
	All3	-1.171***	(0.170)	-0.561***	(0.161)	0.098	(0.092)
Constant		3.683***	(0.114)	3.302***	(0.104)	2.282***	(0.065)
R-squared		0.063		0.033		0.004	
N		2,170		2,210		2,210	

Standard errors in parentheses, bootstrapped standard errors. *** p=0.001, ** p=.01%, * p=.05.
Note: 40 households moved before the first critical period

ment assigned and treatment received. We also run the instrumental variables regressions to estimate the local average treatment effects, instrumenting for treatment received with treatment assigned. The results are not significantly different than those in Table 5, and are shown in Appendix C.

4.2 Evidence that Long-Run Behavioral Change Confounds Estimates of Price Response

It is clear from the treatment effects that all households exposed to critical prices exhibit demand reductions at those times. What is not clear from analyzing household level treatment effects is whether those demand reductions are a result of short-run response to price or long-run behavioral change.Perhaps portal-assigned customers really logged on to learn about critical pricing that may have only been announced with a couple hours of notice. Alternatively, they might have set a different desired peak temperature upon assignment, recognizing that they may at some point be exposed to high critical prices. As an example of this, Fowlie, Wolfram, Spurlock, Todd, Baylis, and Cappers (2017) find that consumers reduce their consumption on non-critical price event peak days as well, even though prices are lower – the consumers clearly responded to being exposed to critical prices, even if they did not exhibit the expected short term response to any particular hour's realized price. In fact, the same long-run response could occur for the technology treatments as well. It is true that demand reductions are greater on critical days than peak days, but they may arise simply because it takes more energy to reduce temperature a given amount on a (hotter) critical day than a peak day.

In our setting, a long-run response to treatment assignment is evident if we compare critical demand reductions under VPP and TOU pricing. Without any long-term effect of price treatment assignment, we would expect similar demand for VPP and TOU pricing when prices are the same (ignoring substitution between peak periods across days). This is not what we observe. For example, in critical periods (when both price assignments face the same 46c price), TOU-assigned households reduce demand more than VPP-assigned households for the information treatments,. This suggests a spillover effect due to the long term consumer response to the higher average TOU price faced during non-critical peak periods. On the other hand, PCT-enabled households exhibit the same critical demand reductions, regardless of whether they are assigned to VPP or TOU pricing. This raises the question of whether demand reductions in the two information treatments involve any short-run response, or whether they merely reflect long-run responses to a systematically higher peak price. If the latter, there is limited scope for using price variation to reduce critical demand and it raises the question of whether consumers are overly reducing demand during non-critical periods. In the next section, we estimate the short-run price elasticities to evaluate this further.

5 Demand Estimation

As previously discussed, the above treatment effects describe the value of these treatments for generating long-run conservation, but they do not directly speak to the ability to align supply and demand in response to short-run shocks: a fundamental challenge in utility markets such as electricity and water that is the precise reason why these dynamic pricing experiments were designed. The optimal experiment to assess short-term price response during peak periods when using different technologies would be to first randomly assign households to the control and technology treatments. Then the utility would randomly vary the realized prices (across households) in each peak period. The experiment we study reflects what has previously been done by utilities and in the academic literature (with the exception of having a control group with constant price). Instead of varying price at the household X time level, variation in price across households at any point in time instead reflects household random assignment to a price treatment, which the household remains in for the course of the experiment.

The estimated treatment effects thus measure long-run price elasticities because they evaluate how average demand over the summer varied in response to assignment to different long run price processes. While such long-run demand elasticities are informative for many managerial and policy questions, they do not answer the question of whether the short-run dynamic price variation is creating corresponding short-run changes in demand. If the risk of exposure to a large temporary price shock generates long-run, instead of short-run, response, the demand response will be reduced because of the surplus lost from having demand reductions in periods without the price increase. This will either not achieve the temporary demand reduction goals or will require a disproportionately larger temporary price increase.

In this section we begin by describing confounds that can arise when estimating short-run price response using variation that is only exogenous because of permanent cross-household random assignment. Next, we define how the randomly assigned control group can be used to conduct a demand analysis that conditions on the unobserved states that shift demand and would otherwise create a price endogeneity problem. Demand estimates using this approach illustrate that only automated demand response technology (i.e. the PCT) achieves non-negligible response to temporary price variation.

5.1 Short-Run Elasticity Confounds from Household Random Assignment

The experiment we analyze, like most dynamic pricing experiments conducted by electric utilities, uses a randomized control trial in which households are randomly assigned to a pricing plan that either includes, or not, price variation that endogenously responds to current demand and supply conditions. Interestingly, the within-household, short-term price variation includes no randomization. There is some exogeneity in the prices faced by different customers at any point in time because of the initial assignment to dynamic pricing or not, but response to that variation may be confounded with permanent reactions to the initial household random assignment.

The following example illustrates how permanent reactions to household random assignment can confound inferences of short-run price elasticities. Suppose a customer assigned to a dynamic pricing plan (TOU or VPP), immediately reacts to assignment by setting a higher maximum internal temperature on her thermostat, while there is no corresponding reaction by a customer assigned to a control group with no dynamic pricing. When price increases because of a high external temperature, the household which set the higher desired internal temperature will use less electricity than the control household, even if the treated household is unaware of a price change. This creates the illusion of a short-run price response, when in fact there was none. The response was generated by a change in demand conditions that are correlated with price, but would occur even if price did not vary.

To illustrate where this bias arises econometrically, we can define a linear demand regression extending equation (1) to consider short-run, time-specific, realizations of the outcomes:

Y_{it} denoting electricity consumption by household i in a given hour t. Adding in the price faced by the household during that time period, P_{it}, and fixed effects for each household, ζ_i, and time period, η_t, we get the following panel data regression:

$$Y_{it} \quad = \beta_a P_{at} + \alpha_0 + \alpha_1 \tilde{A}_i + \eta_t + \zeta_i + \epsilon_{it}. \tag{2}$$

The β_a coefficients measure price response for each treatment group a. The time period fixed effect, η_t, shifts inference to cross-individual price variation within a given time period. Such variation only exists because of the random assignment suggesting a plausible source of exogenous price variation. The identifying assumption of such an approach can be made clear by considering a treatment-time fixed effect, η_{at}, which after conditioning upon it would not leave any remaining price variation to identify the βs.

While it is implausible that treatment effects vary across all time periods t, it is plausible that the treatment effects may systematically differ across certain types of days, e.g. hot days vs. cold days. Since utilities should endogenously vary price by such factors, that would invalidate the identifying assumptions of equation (2) (i.e., if there is a treatment-specific time effect, η_{at}, then $\eta_{at} - \eta_t$ is included in the ϵ_{it} in equation (2), which would then be correlated with P_{at}). The econometrician might try to introduce observable differences across periods and interact them with treatment to reduce bias, but unobserved factors shifting demand differently for treatment and control groups still create problems for identification.

5.2 Conditioning on Unobserved Shocks

Our approach to identification is to focus inference on the price variation that occurs conditional on a given realization of η_{at}. We do not observe η_{at}, but the demand by control households should be driven by the same underlying aggregate utility shocks (e.g., both experience the same high or low temperatures) that affect prices. Letting ξ_t represent the true underlying shock to utility, this means that $\eta_{at} = \eta_a(\xi_t)$. Thus at any point in time, the expected demand intercept in this linearized equation is $\gamma_a(\xi_t) = \alpha_0 + \alpha_1 \tilde{A} + \eta_a(\xi_t)$. The conditional expectation function for the demand for a given treatment, a, and the control group which does not face price variation is:

$$E[Y_{it}|A = a] = \beta_a P_{at} + \gamma_a(\xi_t), \tag{3}$$
$$E[Y_{it}|A = 0] = \gamma_0(\xi_t).$$

With a fixed control group price, the unobserved shock is absorbed into the control group usage, $\gamma_0(\xi_t)$. While this shifts the function, we can still reasonably assume γ_0 is strictly monotonic (and thus invertible). We can then use the latter equation and define the inverse function $\xi_t = \gamma_0^{-1}(E[Y_{it}|A = 0])$ to rewrite the first conditional expectation function as:

$$E[Y_{it}|A = a] = \beta_a P_{at} + \gamma_a\left(\gamma_0^{-1}(E[Y_{it}|A = 0])\right). \tag{4}$$

This control function approach allows for the estimation of $E[Y_{it}|A = a]$ where a time treatment fixed effect, γ_{at}, is substituted with a flexible function of $E[Y_{it}|A = 0]$.[11]

Table 6 reports a series of regression specifications that illustrate the value of the control function approach. The first two specifications include an ordinary least squares regression and a specification that further adds meter ID fixed effects. Both find a positive price coefficient for all technology treatments because they ignore that the VPP price level is higher on high demand days. Next, we add technology-time fixed effects, where time is date-hour, e.g. 2-3pm on June 10th. This is as close as possible to estimating γ_{at}, where we pool the different pricing plan assignments (TOU and VPP) within each set of technology fixed

[11] One critique of control function approaches (which can also be levied against our nonparametric approach of conditioning on control group demand) is that they rely on an assumption that all unobservables can be reduced to a scalar index of the unobserved demand shock. While econometric models are typically written this way (i.e. appending a single unobservable to the outcome equation), in reality there may be multiple unobservables that differentially affect treatment and control groups. We discuss the implications in Appendix D.

effects (recall that a fixed effect at the technology-pricing-time level would leave no remaining price variation). This focuses inference of price response on cross-household variation in price (for households with the same technology assignment, within any given hour of the summer) that arises from the initial random assignment to either TOU or VPP. The baseline price coefficient (for the portal) is now downward sloping but still statistically insignificant. The Price*IHD coefficient is large enough that the demand curve would still not be downward sloping, even ignoring the insignificance. Demand from PCT and All3 treatments exhibit much greater downward sloping demand curves that are significantly more negative than the baseline portal by -1.449 and -1.617 respectively.

The fourth set of results substitutes the control function approach described above for the time fixed effects. The control function includes $E\left[Y_{it}|A=0\right]$ linearly as well as taken to the 2nd, 3rd, and 4th exponents. Each of these are further interacted with the treatment (technology and price) to allow for a flexible control function. Due to the large number of interactions, we do not report these coefficients. The control function does in fact control for the hour specific demand shocks such that coefficients are still negative as in the fixed effects case. But the control function flexibility which allows each treatment to behave differently for a given level of control group demand yields substantial changes in some of the price coefficients. The baseline (Portal) price coefficient is more negative and now statistically significant. The The Price*IHD coefficient is still insignificant. On the other hand, the price coefficient for the PCT and All3 are -2.218 and -2.500 relative to the portal which is a greater than 50% change from the fixed effect specification. This highlights the value of the control function approach in isolating the exogenous short-run price variation such that we would otherwise under-estimate price responsiveness by not adequately removing unobservables that correlate price and quantity.

The final specification estimates the same model, but restricts the data to households on the VPP plan. The logic for this is that the presence of the TOU households only provides inference by comparing them with VPP households on any given day. The resulting VPP-only specification focuses only on within household price variation and finds similar estimates. The similarity suggests that the control function is adequately accounting for how the long-run TOU vs. VPP responses can manifest in different treatment effects under different market conditions.

Table 7 reports the same set of results but using a log-log model where both electricity usage and price have been transformed using the natural logarithm. The coefficients in this model have the added benefit of being directly interpreted as elasticities. We observe a similar evolution of results across the specifications. There are however some notable differences. First, the baseline price coefficient/elasticity which represents the portal assignment documents a statistically significant elastcity, but at a modest value of -0.04. The IHD is never statistically different from the portal assignment. The PCT and IHD elasticities are more than 5 times greater than the portal and statistically significant. These are still rather small at roughly -0.2. This suggests that the prices used in the experiment are not high enough from a short-term profit maximization perspective, likely due to the fact that prices in this market are heavily regulated. These elasticities do, however, suggest the firm has the ability to temporarily reduce demand via short-run price changes if consumers are enabled with automation technology.

In summary, the control function approach substantially changes the magnitude and statistical significance of estimated elasticities relative to fixed approaches similar to applications in other work. Demand is more elastic in all of the control function specifications, but only automation provides nontrivial demand elastcities.

5.3 Non-Parametric Estimates of Demand

A criticism of control functions is that the functional form may not reflect the system of equations generating the data (Wooldridge, 2015). For example, the demand equation above may not be linear (or log-linear). Thus our preferred approach to estimate demand is fully non-parametric. We provide an overview of the non-parametric identification arguments in

Table 6: Linear Regressions of Price Response

	OLS	**FE**	**FE**	**CF**	**CF, VPP Only**
Price	3.206***	3.864***	-0.380	-0.486**	-0.455**
	(0.242)	(0.134)	(0.231)	(0.132)	(0.141)
Price*IHD	-0.225	0.050	0.434	-0.115	-0.128
	(0.395)	(0.215)	(0.348)	(0.151)	(0.164)
Price*PCT	-2.693***	-2.118***	-1.449**	-2.218***	-2.384***
	(0.382)	(0.223)	(0.365)	(0.210)	(0.229)
Price*All3	-2.227***	-2.263***	-1.617***	-2.500***	-2.716***
	(0.411)	(0.231)	(0.370)	(0.223)	(0.256)
Fixed Effects	Treatment	Meter	Tech-Hour, Meter	Meter	Meter
R-squared	0.026	0.505	0.655	0.643	0.636
N	777,542	777,542	777,542	777,542	395,557

Standard errors are clustered by meter ID and bootstrapped for control function specifications to accommodate error in the generated regressor. The latter two specifications substitute a control function interacted with technology treatment for the technology-hour-fixed effects. *** p=0.001, ** p=.01%, * p=.05.

Table 7: Log-Log Regressions of Price Response

	OLS	**FE**	**FE**	**CF**	**CF, VPP Only**
LnPrice	0.125***	0.341***	-0.028	-0.040***	-0.037***
	(0.018)	(0.011)	(0.017)	(0.010)	(0.011)
LnPrice*IHD	-0.024	-0.002	0.007	-0.005	-0.005
	(0.019)	(0.017)	(0.025)	(0.013)	(0.013)
LnPrice*PCT	0.103***	-0.187***	-0.145***	-0.207***	-0.207***
	(0.023)	(0.019)	(0.027)	(0.018)	(0.018)
LnPrice*All3	0.078***	-0.168***	-0.122***	-0.212***	-0.214***
	(0.024)	(0.020)	(0.028)	(0.020)	(0.020)
Fixed Effects	Treatment	Meter	Tech-Hour, Meter	Meter	Meter
R-squared	0.020	0.501	0.645	0.633	0.618
N	775,953	775,953	775,953	775,953	394,356

Standard errors are clustered by meter ID and bootstrapped for control function specifications to accommodate error in the generated regressor.The latter two specifications substitute a control function interacted with technology treatment for the technology-hour-fixed effects. *** p=0.001, ** p=.01%, * p=.05.

Appendix G. In practice, we calculate the conditional expectation of demand at each observed price point and state, ξ_t and then evaluate whether observed price changes are generating observable changes in demand. We therefore estimate $E\left[Y_{it}|A=a, P_t=p, \widehat{\gamma_{0t}}\right]$, where $\widehat{\gamma_{0t}} = E\left[Y_{it}|A=0\right]$. We therefore estimate $E\left[Y_{it}|A=a, P_t=p, \widehat{\gamma_{0t}}\right]$, where $\widehat{\gamma_{0t}} = E\left[Y_{it}|A=0\right]$.

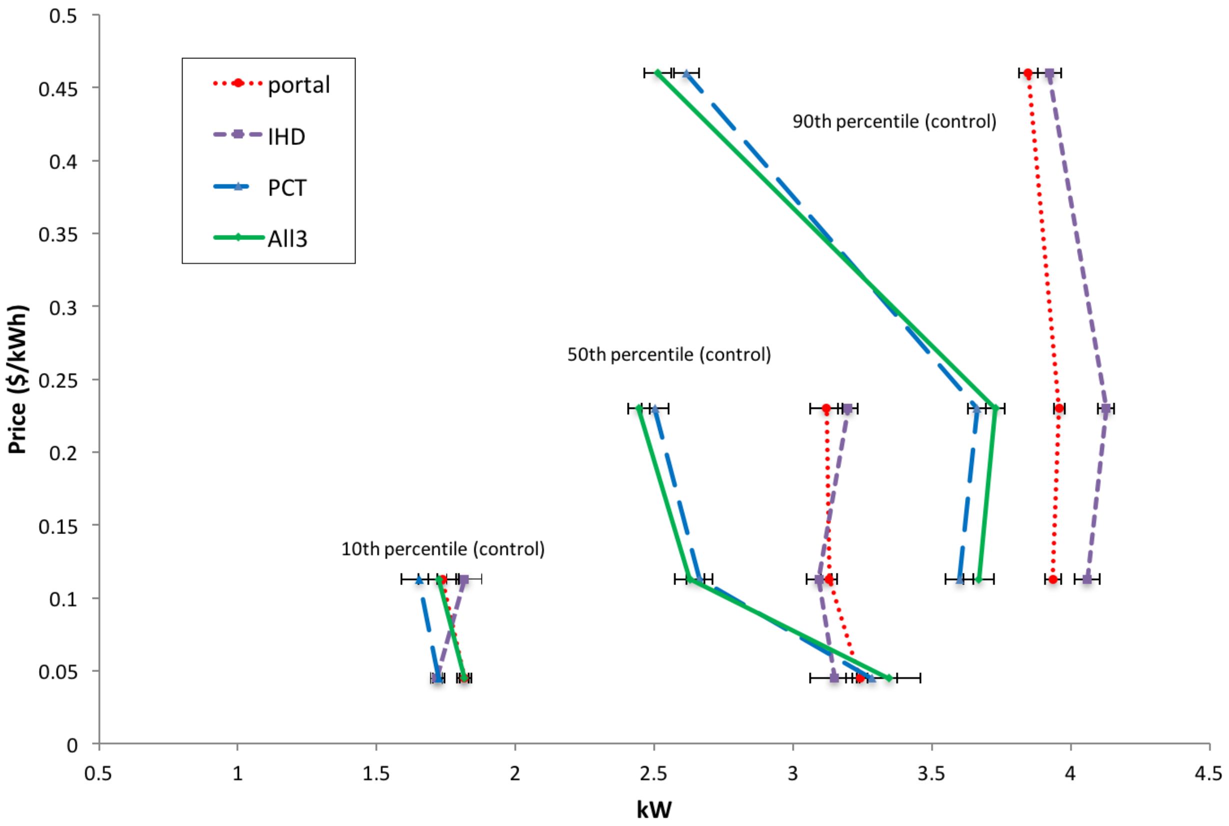
We simply calculate the average consumption for each technology treatment at each price point across those hours in which control group consumption is at a certain percentile. This non-parametric approach avoids any structure in either the first stage or second stage estimates; the only other paper of which we are aware that uses this type of approach is Roberts (2013), who similarly conditions on unobserved auction valuations by conditioning on the reserve prices. In that context, there is an assumed monotonic relationship between the unobserved heterogeneity and reserve prices. In our application, we use the flat pricing in the control group to give us the monotonic relationship that is required.

In practice, to condition on the control group demand $\widehat{\gamma_{0t}}$, we estimate this conditional expectation function at each decile of $\widehat{\gamma_{0t}}$ observed in the data (to see the full set of curves for IHD versus All3 as an example, see Appendix E). Inevitably, we will not observe all prices within all states (e.g. a high electricity price may not be observed on cool summer days), so when we plot these demand estimates, some price points will be missing (typically at the higher or lower price levels).[12] However, we do observe substantial variation in VPP prices conditional on the unobservable (i.e. at each decile of control group demand), indicating that there is still substantial exogenous variation in price that can be used in estimation.

Figure 7 depicts the estimated demand and price points (with lines denoting potential curves) for the 10th, median, and 90th percentiles of control demand during peak periods for all four technology treatments. The All3 treatment is indicated with the green solid line, the PCT with the blue long-dashed line, the IHD with the purple short-dashed line, and the portal with the red dotted line. The 95th percent confidence intervals for each

[12] It is important to note that the percentiles are not a percentile across consumers within the control group, but rather a percentile across time for the entire control group. There is no change in consumer composition.

Figure 7: All3 vs. IHD Demand: 10th, 50th and 90th Percentiles of Control Demand



point on the demand curves are depicted with error bars. The need to condition on the unobserved shock is evident in the graph since the demand curves shift upwards and to the right as we move to higher percentiles of control group usage. As noted earlier, the home automation included in the PCT and All3 demand curves clearly increases demand response. The demand responses we observe at each decile of control demand are summarized in Table 8 for the four technology treatments, measured in the kW adjustment per 10 cent increase in price.

On median usage days by the control group, the largest demand response occurs between 4.5 cents and 11.3 cents, because these days are presumably not too hot and thus a set of consumers are willing to have their PCTs shut off, given the comfort/savings tradeoff. A ten cent price increase reduces consumption by 1.361 kW for PCT and 1.183 kW for All3, as compared to -0.014 kW and -0.147 kW in the 11.3 to 23 cent range. In contrast, on 90th percentile days (which would likely be very hot), the demand responses are much smaller for the PCT and All3 treatments, presumably because too much comfort is sacrificed for the savings.

That said, even on these higher usage days, the automated response is much bigger than for the information technologies: portal and IHD. These technologies have insignificant price responses for the entire range of prices, except the lowest price range around median usage and the highest price range on 80th and 90th percentile control demand days (for the latter, a ten cent price increase lowers consumption by between 0.027 kW and -0.060 kW). All of the PCT and All3 demand responses are significant and substantially larger.

These estimates conditioned on each price point and control group demand level illustrate the importance of avoiding the linear or log-log demand functional form assumptions used above. Clearly, the slope and elasticity vary across price points and control group demand conditions. The estimates in Table 6 (falsely) assumes a linear relationship between usage and with price (or log usage and price) that is invariant to the level of control demand. The non-linear relationships shown in the demand curves in Figure 7 actually have intuitive

Table 8: Demand Response to Price (kWh/hr reduction for each $.10 price increase)

		Portal			IHD		
	Price:	4.5-11.3	11.3-23	23-46	4.5-11.3	11.3-23	23-46
	10	-0.116			0.089		
	30	0.090	-0.088		-0.046	-0.118	
	40	-0.059	-0.032		-0.251***	-0.043	
Control	50	-0.164**	-0.008		-0.243***	0.051	
Usage	60	-0.208**	0.033		-0.368***	0.045	
	70	-0.045	-0.010	-0.048	0.146	-0.025	-0.094
	80		0.026	-0.027		0.050	-0.060*
	90		0.019	-0.048***		0.001	-0.051**

		PCT			All3		
	Price:	4.5-11.3	11.3-23	23-46	4.5-11.3	11.3-23	23-46
	10	-0.341***			-0.383***		
	30	-0.325***	-0.522***		-0.330**	-0.533***	
	40	-1.323**	-0.002**		-1.280***	-0.126*	
Control	50	-1.361***	-0.014		-1.183***	-0.147*	
Usage	60	-1.293**	-0.200***		-1.188**	-0.247***	
	70	-0.859***	-0.246***	-0.345***	-0.772***	-0.326***	-0.331***
	80		-0.212***	-0.105**		-0.173***	-0.193***
	90		0.149***	-0.261***		-0.094*	-0.358***

*** p=0.001, ** p=.01%, * p=.05. For the 20th percentile of control group usage, we only observe demand at $0.045 and $0.46, so we cannot calculate the intermediate slopes.

implications for firm pricing: there is much greater response on the upper portion of the demand curve on high demand (e.g. really hot) days, but such response is achievable at much lower prices on median demand days.

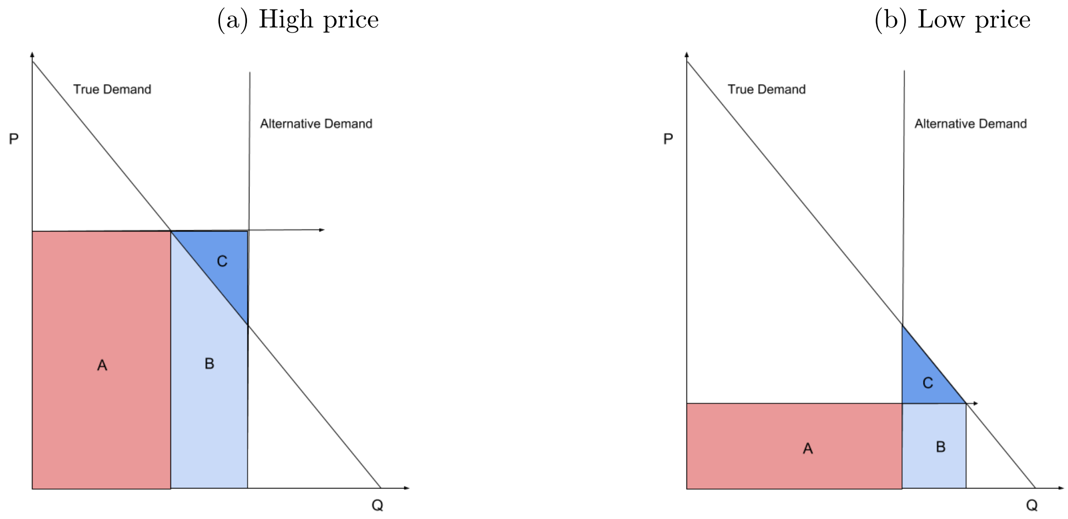
5.4 Consumer Surplus from Automated Demand Response

The primary role of automation technology is to eliminate the demand inelasticity that arises from consumers' costs of adjusting demand to changing prices. The PCT allows consumers to provide a preference input of comfort vs. savings into a demand algorithm which should improve short-run demand elasticity, as exhibited in the demand curves in Figure 7. To the extent that demand input and any short-run inputs can replicate their true preferences, one can approximate the welfare gain from the automation technology by comparing the All3 demand curve, which has both automation and information available, to the IHD demand curve. This estimate is not a precise welfare calculation for a variety of reasons including: i) the comfort vs. savings input likely cannot define the complex policy function that would underly the consumer's optimal choice under each state, ii) there are costs of adjusting the demand in the short run, the savings of which are welfare improvements of the automation technology; iii) these costs may in practice lead to individual level discontinuities in the demand function, iv) there are externalities from demand reduction (e.g. reduced reliance on more polluting technologies), etc., and v) we do not consider supply-side response in demand side welfare or quantify the supply side welfare effects because we do not observe the relevant costs or supply functions. However, under some reasonable assumptions and approximations using piece-wise linearity, we can approximate consumer benefits from adding automation technology (a PCT).

Figures 8a and 8b illustrate the demand-side tradeoffs that occur when consumers respond to short-run price variation with long-run response. Consider a consumer with a "True Demand" that is downward-sloping in the short-run, but, due to costs of adapting to short-run price variation, responds in the short-run with a perfectly inelastic "Alternative Demand". If the consumer faces either a high critical peak price (8a) or a low price (8b), we would expect Alternative Demand to intersect True Demand at an intermediate price level the consumer may never actually face. Under the high price, the consumer overspends by the area B+C in Figure 8a, and realizes a surplus loss of C. When the price is low, the consumer underspends and realizes a surplus loss of C in Figure 8b. The upfront setting of inelastic Alternative Demand will set the size of these two triangles based on the frequency with which each price is expected. Note that we are ignoring short-term demand shocks to the consumer in this example. The consumer surplus lost from the inability to respond to short-run price variation, or alternatively the value of a technology that could replicate true demand, would be the expectation of the area C.

We do not observe True Demand to estimate these effects, but under some reasonable

Figure 8: Consumer Surplus Lost with Adjustment Frictions



assumptions we can approximate the consumer surplus gains from automation technology. Specifically, we assume that the closest treatment to True Demand is the All3 which allows them to program their thermostats accounting for their price elasticity (PCT) and have the increased salience and information associated with the IHD. Recognizing that the All3 demand is an imperfect representation of the True Demand, we first quantify the surplus loss from using an All3 which does not yield the True Demand. Next, we show that the surplus loss when using only an IHD is that same loss plus a quantifiable surplus difference between the IHD and All3 treatment.

Suppose that the only difference between the All3 demand curve and the true demand curve results from a short-term, stochastic demand shock, ϵ_t , as depicted in Figure 9a. The idea is that ϵ_t shifts the true demand to the left or right in ways that the consumer does not respond to in the short-run, e.g. the consumer may be in or out of the house more on a given day implying an increased or decreased need for electricity to cool the home, but does not provide a short-run adjustment to the automation technology. (Note that this is exactly the inelasticity a Nest can adapt to because of motion sensors.) The lost welfare from the demand shock, ϵ_t, in the All3 treatment group is equal to $\frac{1}{2}p'\epsilon^2$, where surplus is assumed to be represented by a triangle with constant slope p' such that the height of the triangle is equal to the base multiplied by the slope. This is the light triangle (bcf) for a negative shock and the dark triangle (acd) for a positive shock. Mathematically, we can represent this welfare left on the table when a PCT-IHD combination does not fully reflect true demand as:

$$W^{*} - W^{All3} = \frac{1}{2}p' \int_{-\infty}^{\infty} \epsilon_t^2 f(\epsilon_t) d\epsilon_t. \tag{5}$$

When a consumer is only equipped with an IHD to execute demand and exhibits the inelasticity shown in the previous section, the surplus changes depicted in Figure 9b are realized. The same demand shocks depicted in the previous figure can occur, with negative, zero or positive values of ϵ_t, respectively representing the three downward sloping demand curves from left to right. Executed demand is, however, perfectly inelastic as represented by the IHD. With a positive demand shock, the welfare lost in the IHD treatment is equal to the dark blue triangle (deh), i.e. demand is moved closer to the over-consuming IHD demand. With a negative shock of the same magnitude, it is equal to the triangle (bej) with a greater surplus loss arising from the increased gap between true demand and the overconsumption under IHD. Mathematically, we can represent this welfare left on the table when a consumer only has access to an IHD to learn and help respond to prices according to her true demand as:

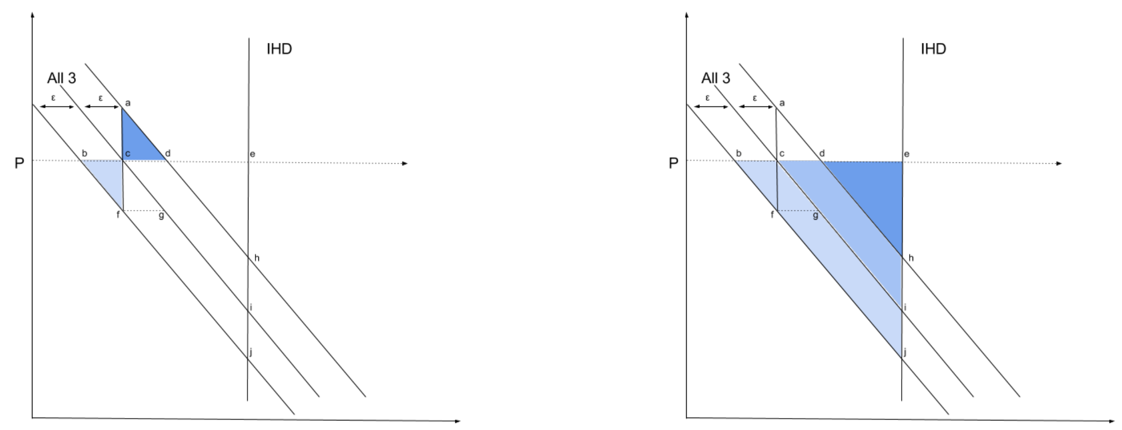
$$W^{*} - W^{IHD} = \int_{-\infty}^{\infty} \frac{1}{2}p' \left(\Delta q + \epsilon_t\right)^2 f(\epsilon_t) d\epsilon_t = \frac{1}{2}p'\Delta q^2 + \frac{1}{2}p' \int_{-\infty}^{\infty} \epsilon_t^2 f(\epsilon_t) d\epsilon_t$$

in which we define $\Delta q \equiv q^{*}(p) - q^{IHD}(p)$. At the price P in the figure, Δq is negative, and so a negative ϵ leads to an even greater loss in welfare. It should be noted that the $p'\Delta q\epsilon$ term drops out of the expression since we assume a symmetric distribution for the ϵ around the true demand curve, and so the negative ϵ are offset by the positive ones in this term.

The value of adding a PCT to an IHD to create the All3 treatment can therefore be

Figure 9: Consumer Surplus Lost Due to Differences in Consumption Relative to True Demand Curve

(a) All3 welfare lost (b) IHD welfare lost



represented as:

$$\Delta W^{PCT} = W^{All3} - W^{IHD} = \frac{1}{2}p'\Delta q^2.$$

Note that the value of a PCT over a portal can be similarly calculated and would be a comparable magnitude based on the demand curves represented in Figure 7. Since the relative value does not depend on the distribution of the demand shocks (which affect both surplus calculations by the same amount and thus cancel out in the difference), we can estimate this approximate consumer surplus change using the estimated piece-wise linear demand curves shown in Figure 7.

To estimate the total effect on consumer surplus, we must integrate over all states. We do this by considering each decile of the control demand, solving for the surplus effect at each price observed in that decile, and then calculating the weighted average based on how frequently that price was charged within that decile of control demand. We calculate the welfare over the entire summer of 2011 by integrating over the realized states. Assuming piece-wise linearity as described above, we estimate the per-household welfare gain from adding automation to an IHD is \$29 across the entire summer of 2011, or \$560 NPV with a 0.95 discount rate (and assuming a similar reduction in later summers). This shows that the average surplus gain from the PCT outweighs the cost. However, the cost of advertising to communicate the benefits could make this unprofitable, and consumer adoptions may be unlikely considering the literature on the energy efficiency gap (Allcott and Greenstone, 2012).

One caveat to these calculations is that they depend on the user being able to reflect their average True Demand curve using the PCT, given the ability to input their personalized comfort/cost tradeoff. If this is not the case, then we are overestimating the consumer welfare gain from the PCT. One potential reason for this could be that consumers need to learn how to use the PCT over time, or that they need to learn about their consumption patterns. We test this hypothesis and find no evidence of changing treatment effects over time for any of the technology X price treatment combinations, after controlling for the realized peak prices.

However, consumers are not the only beneficiaries; the utility benefits from being able to reduce consumption during periods of critical peak demand. Our utility partner estimates that each kWh reduction during the peak demand period is valued at \$700 NPV in delayed capacity investment. The full welfare benefits of home automation technologies such as the PCT includes these firm benefits, the consumer surplus we calculated, as well as the environmental benefits from reduced carbon emissions. Given the large utility benefits, it is not surprising that all households in the utility area have now been given PCTs. We expect the firm benefits to increase further with more sophisticated dynamic pricing schedules that truly reflect the marginal costs of production (within the regulatory constraints).

6 Conclusion

This paper contributes to the literatures on information economics and dynamic pricing. Dynamic (or surge) pricing is being used in many new contexts, including ride-sharing (Hall, Kendrick, and Nosko, 2015) and Airbnb rentals (Hill, 2015). In these two-sided platforms, price information is revealed through the platform and the pricing mechanism helps equates demand and supply (Hall, Kendrick, and Nosko, 2015; Chen, Chevalier, Rossi, and Oehlsen, 2017; Cachon, Daniels, and Lobel, 2017).

Despite the benefits of information provision, in many contexts consumers are unable to shift their consumption in response to short-term price changes without incurring large adjustment costs. Electricity consumption is one such context, as are other input goods which are consumed in repeated or continuous, automated processes. We have demonstrated that sophisticated home automation technology that can respond to dynamic pricing can greatly increase price response. In addition to leading to far greater demand reductions than information and communication technologies, the automation technology leads to short-term price elasticity by eliminating adjustment costs. In contrast, households with information or communication technology exhibit a shift in their demand curves, but the demand curves remain inelastic.

In order to estimate these demand curves, we propose an approach that isolates short-term exogenous price variation even when the randomization occurs at the household level instead of the prices themselves. In contexts in which a control group can be assigned flat prices, conditioning on the control group's percentile of usage is akin to conditioning on treatment-specific responses to the unobserved demand shock. This leaves any remaining price variation as exogenous to households in the experiment and therefore viable for estimating short-term price response. Methods which leverage random variation across treatment groups in price schedules actually estimate the long-run response to treatment assignment, rather than the short-term price response of interest. If policymakers are able to induce short-term price response, VPP pricing can then lead to much greater efficiency gains than TOU (Braithwait, Hansen, and O'Sheasy., 2007; Hogan, 2014).

We show that consumers who can automate their response to dynamic pricing are the only ones who respond to actual real time prices, whereas other consumers respond with long-run adjustments to treatment assignment. This paper is one of the first in a bourgeoning literature on automation in marketing. Current working papers are studying the effects of automation in other applications such as sales force pricing decisions (Shichor and Netzer, 2018) and advertising. Choi, Mela, Balseiro, and Leary (2017) review the literature on advertising display markets in which over 78% percent of banner ads were bought using automated, programatic buying in 2017[13] We anticipate a further acceleration of the adoption of automation, given the terrific potential of such technologies in reducing frictions and enabling real-time response.

References

Allcott, H. (2011): "Rethinking Real-Time Eectricity Pricing," *Resource and Energy Economics*, 33, 820–842.

Allcott, H., and M. Greenstone (2012): "Is There an Energy Efficiency Gap?," *The Journal of Economic Perspectives*, 26(1), 3–28.

Borenstein, S., and S. Holland (2005): "On the Efficiency of Competitive Electricity Markets with Time-Invariant Retail Prices," *RAND Journal of Economics*, 36(6), 469–493.

Braithwait, S., D. Hansen, and M. O'Sheasy. (2007): "Retail Electricity Pricing And Rate Design In Evolving Markets," Discussion paper.

Cachon, G. P., K. M. Daniels, and R. Lobel (2017): "The role of surge pricing on a service platform with self-scheduling capacity," *Manufacturing & Service Operations Management*, 19(3), 368–384.

Cappers, P., C. Goldman, and D. Kathan (2010): "Demand Response in U.S. Electricity Markets: Empirical Evidence," *Energy*, 35(4), 1526–1535.

[13]https://www.emarketer.com/Article/eMarketer-Releases-New-Programmatic-Advertising-Estimates/1015682

Chen, M. K., J. A. Chevalier, P. E. Rossi, and E. Oehlsen (2017): "The value of flexible work: Evidence from uber drivers," National Bureau of Economic Research.

Choi, H., C. Mela, S. Balseiro, and A. Leary (2017): "Online Display Advertising Markets: A Literature Review and Future Directions," working paper.

Compiani, G., P. Haile, and M. Sant'Anna (2017): "Common Values, Unobserved Heterogeneity, and Endogenous Entry in U.S. Offshore Oil Lease Auctions," working paper.

Erdem, T., S. Imai, and M. P. Keane (2003): "Brand and Quantity Choice Dynamics Under Price Uncertainty," *Quantitative Marketing and Economics*, 1(1), 5–64.

Faruqui, A., S. Sergici, and A. Sharif (2010): "The impact of informational feedback on energy consumption-A survey of the experimental evidence," *Energy*, 35(4).

Fowlie, M., C. Wolfram, C. A. Spurlock, A. Todd, P. Baylis, and P. Cappers (2017): "Default Effects and Follow-on Behavior: Evidence from an Electricity Pricing Program," working paper.

Guerre, E., I. Perrigne, and Q. Vuong (2009): "Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions," *Econometrica*, 77(4), 1193–1227.

Haile, P. A., H. Hong, and M. Shum (2003): "Nonparametric Tests for Common Values at First-Price Sealed-Bid Auctions," NBER Working Paper No. 10105.

Hall, J., C. Kendrick, and C. Nosko (2015): "The effect of Uber's surge pricing: A case study," The University of Chicago Booth School of Business.

Harding, M., and C. Lamarche (2016): "Empowering Consumers Through Data and Smart Technology: Experimental Evidence on the Consequences of Time-of-Use Electricity Pricing Policies," *Journal of Policy Analysis and Management*, 35(4), 906–931.

Hendel, I., and A. Nevo (2006): “Measuring the Implications of Sales and Consumer Inventory Behavior,” *Econometrica*, 74(6), 1637–1673.

Henley, A., and J. Peirson (1994): “Time-of-use electricity pricing: evidence from a British experiment,” *Economics Letters*, 45(3), 421–426.

Herter, K., P. McAuliffe, and A. Rosenfeld (2007): “An exploratory analysis of California residential customer response to critical peak pricing of electricity,” *Energy*, 32(1), 25–34.

Herter, K., and S. Wayland (2010): “Residential response to critical-peak pricing of electricity: California evidence,” *Energy 35*, 35(4), 1561–1567.

Hill, D. (2015): “How much is your spare room worth?,” *IEEE Spectrum*, 52(9), 32–58.

Hogan, W. W. (2014): “Time-of-Use Rates and Real-Time Prices,” .

Honka, E. (2014): “Quantifying Search and Switching Costs in the US Auto Industry,” *RAND Journal of Economics*, 45(4), 847–884.

Ito, K. (2014): “Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing,” *The American Economic Review*, 104(2), 537–563.

Jessoe, K., and D. Rapson (2014): “Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use,” *The American Economic Review*, 104(4), 1417–1438.

Joskow, P. L., and C. D. Wolfram (2012): “Dynamic Pricing of Electricity,” *American Economic Review: Papers & Proceedings*, 102(3), 381–385.

Liu, P., J. Ni, and Q. Shen (2017): “The Multi-facetted Effects of Rebates: Evidence from Energy Consumption,” working paper.

Mela, C. F., S. Gupta, and D. R. Lehmann (1997): "The Long-Term Impact of Promotion and Advertising on Consumer Brand Choice," *Journal of Marketing Research*, 34(2), 248–261.

Reiss, P. C., and M. W. White (2005): "Household Electricity Demand, Revisited," *Review of Economic Studies*, 72(3), 853–883.

Roberts, J. (2013): "Unobserved Heterogeneity and Reserve Prices in Auctions," *The RAND Journal of Economics*, 44(4), 712–732.

Seiler, S. (2013): "The Impact of Search Costs on Consumer Behavior: A Dynamic Approach," *Quantitative Marketing and Economics*, 11(2), 155–203.

Shichor, Y. K., and O. Netzer (2018): "Automating the B2B Salesperson Pricing Decisions: Can Machines Replace Humans and When?," working paper.

Stigler, G. J. (1961): "The Economics of Information," *Journal of Political Economy*, 69(3), 213–225.

Wooldridge, J. M. (2015): "Control function methods in applied econometrics," *Journal of Human Resources*, 50(2), 420–445.

Appendix A: Randomization Check and Compliance

In this appendix, we check the validity of the randomization by testing whether there are any significant pre-treatment outcome differences between the treatment and control groups. Specifically, we estimate Equation 1 using the 2010 summer electricity usage as the outcome and report the results in the following table. The first two columns consider all days and times, while the second two columns considers just the peak hours on weekdays. There are no significant differences from the control group. The smallest p-value for all days and times is 0.329 for the VPP-PCT treatment. The smallest p-value for the peak is 0.246 for TOU-IHD.

While there are no significant pre-treatment differences in electricity usage, we do find two significant differences in demographics. One difference is that households in the control group are significantly more likely to be families. We have no explanation for this and since there are no usage differences, we suspect this is a random difference. The other difference arises from our focus on households who reported having air conditioning. There were 97 households from the control group for whom the air conditioning variable was never filled out (this air conditioning ambiguity was resolved for all treatment groups). There were clearly both AC and non-AC households within this group because the fraction of controls with air conditioning is about 75% compared to 90% or more for most of the treatment groups. Similarly, only 1.2 % of control households reported not having AC, while 4-9 percent of treatment households reported not having AC. Thus, when we restrict the sample to households observed to have AC, we are dropping some control households that do in fact have AC but did not report it. These households had significantly higher incomes as we see our final AC-only sample yield a control group with significantly lower incomes. These differences do not exist when we conduct our analysis without conditioning on AC.

Table 9: Summer of 2010 kWh Differences Between Treatment Groups and Control

	All Days and Times		**Peak**	
Treatments	**TOU**	**VPP**	**TOU**	**VPP**
Portal	-0.056	0.0056	-0.040	0.065
	(0.085)	(0.088)	(0.116)	(0.118)
IHD	-0.081	-0.006	-0.142	0.006
	(0.092)	(0.092)	(0.122)	(0.125)
PCT	-0.015	0.0935	0.015	0.096
	(0.088)	(0.096)	(0.125)	(0.130)
All3	0.056	0.027	0.052	0.064
	(0.098)	(0.106)	(0.133)	(0.142)
	N=2,178		N=2,178	

Standard errors in parentheses

Our average treatment effects are comparable in both cases and the pre-treatment electricity usage yields no significant differences in either case. We choose to report results conditioning on AC because this is an important variable in the matching approach to recover individual estimates. We do not match to obtain individual effects for non-AC households because the sample sizes are too small. This does however imply that some of the best matches for control outcomes might be missing. If we felt this were a concern we could include those controls that did not fill out the AC variable and the matching would implicitly reveal the AC variable through the pre-treatment usage data.

We also report the differences in critical, peak, and off-peak consumption using treatments assigned rather than treatments received. The results are shown in Table 10 and are almost identical to those shown in Table 4.

Table 10: Hourly Usage by Technology Assigned

(a) Critical Period Hourly Demand by Technology Assigned

control	price	portal	IHD	PCT	All3
3.692 (2.276) N=311	TOU	3.360† (2.114) N=246	3.172** (1.968) N=219	2.597*** (1.971) N=252	2.829*** (2.034) N=205
	VPP	3.461 (2.212) N=271	3.325* (1.982) N=213	2.624*** (2.026) N=248	2.695*** (2.062) N=205

(b) Non-Critical Peak Period Hourly Demand by Technology Assigned

control	price	portal	IHD	PCT	All3
3.339 (2.249) N=317	TOU	3.091 (2.060) N=249	2.969* (1.935) N=223	2.482*** (1.941) N=255	2.718*** (2.032) N=209
	VPP	3.234 (2.223) N=275	3.100 (1.959) N=224	2.735*** (2.016) N=249	2.860* (2.053) N=209

(c) Non-Critical Off-Peak Period Hourly Demand by Technology Assigned

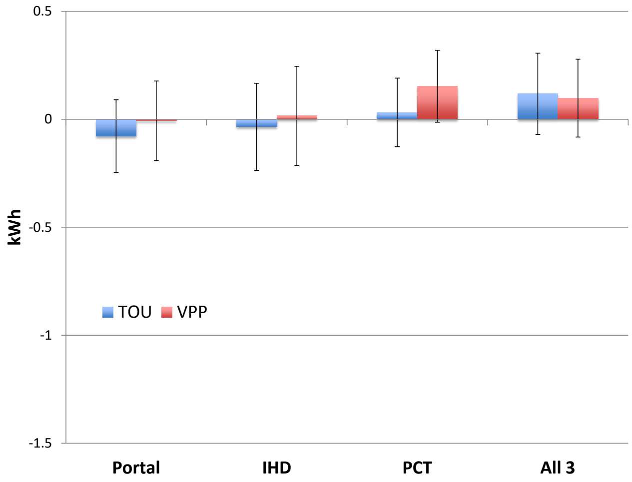
control	price	portal	IHD	PCT	All3
2.303 (1.902) N=317	TOU	2.269 (1.785) N=249	2.231 (1.742) N=223	2.242 (1.733) N=255	2.446 (1.906) N=209
	VPP	2.366 (1.930) N=275	2.294 (1.787) N=224	2.385 (1.873) N=249	2.419 (1.893) N=209

Standard deviation in parentheses. Stars indicate significant differences in means relative to the control group.
*** p=0.001, ** p=.01%, * p=.05. †p=.1

Appendix B: Treatment Effects

Off-Peak Estimates

Figure 10: Off-Peak Demand Reductions by Treatment



Error bars indicate 95th percent confidence intervals. Differences are not statistically significant.

Regressions Using Hourly Data

As a robustness check, we also estimate a regression using the uncollapsed hourly data to identify the treatment effects.:

$$y_{it} = \alpha_0 + \alpha_1 \left(A_i \tilde{\otimes} P_t \right) + \eta_t + \xi_i + \epsilon_{ita} \tag{6}$$

in which P_t is a vector of dummy variables indicating whether the period is a critical, peak, or off-peak time, $\otimes$ is the Kronicker product; α_1 in now a vector which measures the effect of each treatment for each period (critical, peak, and off-peak). We include different combinations of period, hour-in data, and household fixed effects as controls (to include the household fixed effects, we need to include the 2010 usage data). Although these controls are not needed per-se, they help explain variation in usage across the hours in the data, and across households. The household fixed effects also control for any differences across the treatment groups if there was any issue with the randomization. To incorporate them, we include 2010 summer consumption data as well which allows for a differences-in-differences estimation approach. The estimates are largely unchanged across specifications.

Table 11: Treatment Effect Regression Results with Controls

Treatments			(2011 1) estimate	s.e.	(2011 2) estimate	s.e.	(2010 & 2011) estimate	s.e.
Critical	**TOU**	**Portal**	-0.392***	(0.042)	-0.395**	(0.145)	-0.353***	(0.086)
		IHD	-0.478***	(0.046)	-0.477**	(0.153)	-0.487***	(0.091)
		PCT	-1.164***	(0.047)	-1.161***	(0.148)	-1.130***	(0.130)
		All3	-0.971***	(0.049)	-0.972***	(0.156)	-1.087***	(0.133)
	VPP	**Portal**	-0.286***	(0.042)	-0.287	(0.147)	-0.348***	(0.084)
		IHD	-0.337***	(0.046)	-0.335*	(0.150)	-0.334***	(0.089)
		PCT	-1.115***	(0.047)	-1.114***	(0.148)	-1.213***	(0.135)
		All3	-1.152***	(0.050)	-1.154***	(0.160)	-1.203***	(0.132)
Peak	**TOU**	**Portal**	-0.293***	(0.007)	-0.295*	(0.131)	-0.254***	(0.071)
		IHD	-0.336***	(0.008)	-0.349*	(0.140)	-0.339***	(0.074)
		PCT	-0.896***	(0.008)	-0.930***	(0.134)	-0.874***	(0.081)
		All3	-0.722***	(0.009)	-0.755***	(0.143)	-0.836***	(0.085)
	VPP	**Portal**	-0.159***	(0.007)	-0.167	(0.134)	-0.212**	(0.071)
		IHD	-0.202***	(0.008)	-0.208	(0.136)	-0.204**	(0.073)
		PCT	-0.604***	(0.008)	-0.641***	(0.138)	-0.695***	(0.086)
		All3	-0.574***	(0.009)	-0.605***	(0.148)	-0.619***	(0.086)
Off-Peak	**TOU**	**Portal**	-0.095***	(0.003)	-0.097	(0.088)	-0.056	(0.040)
		IHD	-0.034***	(0.003)	-0.031	(0.100)	-0.037	(0.045)
		PCT	0.007***	(0.003)	0.021	(0.092)	0.028	(0.040)
		All3	0.136***	(0.004)	0.14	(0.106)	0.023	(0.045)
	VPP	**Portal**	0.013***	(0.004)	0.013	(0.092)	-0.040	(0.044)
		IHD	0.019***	(0.003)	0.018	(0.096)	0.017	(0.046)
		PCT	0.153***	(0.003)	0.144	(0.100)	0.061	(0.052)
		All3	0.100***	(0.004)	0.1	(0.110)	0.055	(0.043)
peak period			0.354***	(0.030)				
critical period			1.389***	(0.030)				
Hour in data-dummies			N		Y		Y	
Household fixed Effects			N		N		Y	
R-squared			0.019		0.277		0.394	
N			6,285,960		6,285,960		12,535,790	

Standard errors in parentheses, clustered at HH level. *** p=0.001, ** p=.01%, * p=.05.
Third column includes clustering at HH and hour-in-data, includes 6,249,830 hourly observations from the summer of 2010.

Appendix C: LATE Effects Using Instrumental Variables Regression

Table 12: Treatment Effect Regression Results, Collapsed Data

		Critical		Peak		Off-Peak	
Treatments		**estimate**	**s.e.**	**estimate**	**s.e.**	**estimate**	**s.e.**
TOU	**Portal**	-0.342*	(0.149)	-0.215	(0.121)	-0.019	(0.096)
	IHD	-0.567***	(0.160)	-0.361*	(0.143)	-0.080	(0.088)
	PCT	-1.1226***	(0.168)	-0.932***	(0.109)	-0.042	(0.089)
	All3	-0.945***	(0.169)	-0.666***	(0.106)	0.161	(0.120)
VPP	**Portal**	-0.255	(0.153)	-0.103	(0.154)	0.056	(0.099)
	IHD	-0.364*	(0.154)	-0.243	(0.130)	-0.023	(0.110)
	PCT	-1.258***	(0.150)	-0.679***	(0.133)	0.075	(0.109)
	All3	-1.133***	(0.173)	-0.526***	(0.167)	0.123	(0.122)
Constant		3.683***	(0.100)	3.302***	(0.092)	2.282***	(0.071)
R-squared		0.062		0.032		0.002	
N		2,170		2,210		2,210	

Standard errors in parentheses, bootstrapped standard errors
Note: 40 households moved before the first critical period

Appendix D: Two Unobservables in a Control Function

Econometric models typically rely on the assumption that all unobservables can be reduced to a scalar index of the unobserved demand shock (i.e. appending a single unobservable to the outcome equation). In reality there may be multiple unobservables that differentially affect treatment and control groups.

We illustrate how this problem might arise in our context in Figure 11 where the unobservables are two different temperature levels: T_1 and T_2. In the left panel, a given level of the control group demand is represented by the red isoquant labeled $Q(0, P_0)$. The isoquant is concave because reductions in one temperature, e.g. T_1, have increasingly negligible impacts on demand as it approaches 72 degrees such that almost all electricity usage is driven by the higher temperature T_2 as we approach the left side of the graph. The key feature creating an endogeneity problem within the control function approach is that despite control group demand being constant, the different temperature combinations along the red curve could involve different prices. For example, electricity prices may be more reflective of the maximum temperature within a given day such that price is increasing as we rotate away from the diagonal.[14] Such price changes are not exogenous and not what we would like to use for econometric inference. To analyze how this impacts our estimation, we must consider the role of price within these diagrams.

[14] In our data, the high temperature in a day can explain the incidence of the maximum \$0.46 price whereas the low temperature on a day appears to have no explanatory power for those prices. On the other hand, the low temperature has much more explanatory power for the temperature variation between \$0.045 and \$0.23.

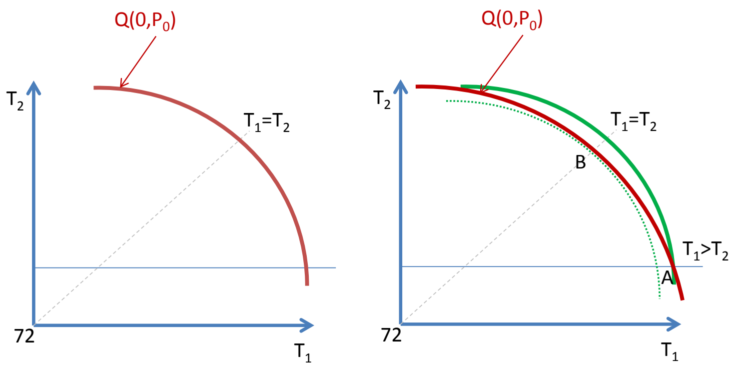


Figure 11: Potential Bias Using Control Functions

The higher are prices and/or the greater is the price sensitivity, the more concave will be the isoquants. In the right panel, the same control isoquant is in red and a more price responsive treatment isoquant is green and the two intersect at the point where $T_1 > T_2$. An increase in T_2 should increase the quantity demanded for both, but to a lesser degree if price is high and influential on demand. Thus, the green isoquant is closer to a hypothetical vertical line that would reflect a pure T_2 increase from the point where $T_1 > T_2$.

To predict the bias, consider how price may vary along the control group isoquant. Conditional on $Q(0, P_0)$, the mix of unobservable temperatures may be at point A or B, where $P(A) > P(B)$ because higher temperatures increase the potential for peak load problems. At A, both the control and treatment groups are on the same isoquant (though demand may still be different). As price drops when moving to B, the treatment group demand would move to the dashed green isoquant which is a lower demand level. The prediction is therefore a positive relationship between price and quantity when price moves because of temperature variation that exists conditional on control group demand.

In our study, all technology-price treatment combinations involve some higher marginal electricity prices but then vary in the technological assistance in adapting to those price changes. The preceding suggests that technologies that have an increased effect on price responsiveness, will have a greater inelasticity bias inferred by control function approaches. To evaluate the potential of these biases, we consider estimation of demand response with and without conditioning on multiple temperature levels that may affect demand (as in the isoquant illustrations). Temperature variation within the hourly observations we analyze may be minimal, so we exacerbate the problem by aggregating our analysis to the entire peak time period during each day. We then use the high and low temperature observed each day to represent the two potential unobservables that could both i) yield identical control group demands for different temperature mixes and ii) still have correlation between price and the particular temperature mix realized.

We therefore reconstruct our estimates from Table 7 using an aggregation of peak demand throughout the day. We then conduct our analysis as in the text of the paper, without accounting for a high and low temperature, and then do it again using the observed high and low temperatures. Column "CF" in Table 13 replicates column "CF" in Table 7 using the daily aggregated data. In the daily aggregation, we find the elasticity for the portal to be -0.032 and statistically significant whereas it was -0.026 and insignificant in the hourly analysis. The daily analysis should be the less reliable as there is more variability in unobservables that could be confounding at the day than the hour level (e.g. more temperature variation within the 2-7pm peak period as opposed to any given hour within that frame). In column "Temp CF" in Table 13 we condition on temperature in both the first stage estimate of the control function as well as the second stage estimate of the demand response of the treatment groups. We see the elasticity for the portal option drops back to -0.013. The IHD elasticity is statistically indistinguishable from the portal in all specifications. When the PCT is added, demand is slightly more elastic in "Temp CF" than "CF". When we restrict the analysis to only variation in the VPP price, which is the preferred specification, both the portal and IHD lack statistical and economic significance in the elasticity estimates, whether controlling for temperature or not. The PCT and All3 treatments exhibit slightly greater elasticity after controlling for temperature which is consistent with the intuition described above (i.e. price responsive treatments will exhibit a bias toward inelasticity when multiple unobserved temperatures may be correlated with both price and demand).

Table 13: Log-Log Regressions of Price Response with Temperature Included in Control Function Estimate

	CF	Temp CF	CF, VPP Only	Temp CF, VPP Only
LnPrice	-0.032***	-0.013	-0.001	-0.014
	(0.010)	(0.014)	(0.010)	(0.013)
LnPrice*IHD	-0.022	-0.011	-0.019	-0.009
	(0.015)	(0.020)	(0.015)	(0.021)
LnPrice*PCT	-0.135***	-0.166***	-0.127***	-0.164***
	(0.018)	(0.022)	(0.019)	(0.022)
LnPrice*All3	-0.141***	-0.146***	-0.128***	-0.141***
	(0.018)	(0.023)	(0.019)	(0.024)
Fixed Effects	Meter	Meter	Meter	Meter
R-squared	0.747	0.746	0.743	0. 742
N	94,939	94,939	48.250	48,250

Standard errors are clustered by meter ID and bootstrapped for control function specifications to accommodate error in the generated regressor

The latter two specifications substitute a control function interacted with technology treatment for the technology-hour-fixed effects. *** p=0.001, ** p=.01%, * p=.05.

In summary, multiple unobservables can confound the estimation of demand when using control functions, however the implications for our analysis appear minimal. We constructed an example loaded toward biasing our estimates by aggregating demand within the day and explicitly ignoring the high vs. low temperature variation that can occur within that window. Even with this, we find minimal if non-existent bias in the demand estimates. Considering that temperature explains 93% of the variation in demand, it is hard to conceive that there are other mixes of unobservables that would create more substantive problems for our demand analysis.

Appendix E: Demand Curve for Each Decile of Control Usage, IHD versus All3

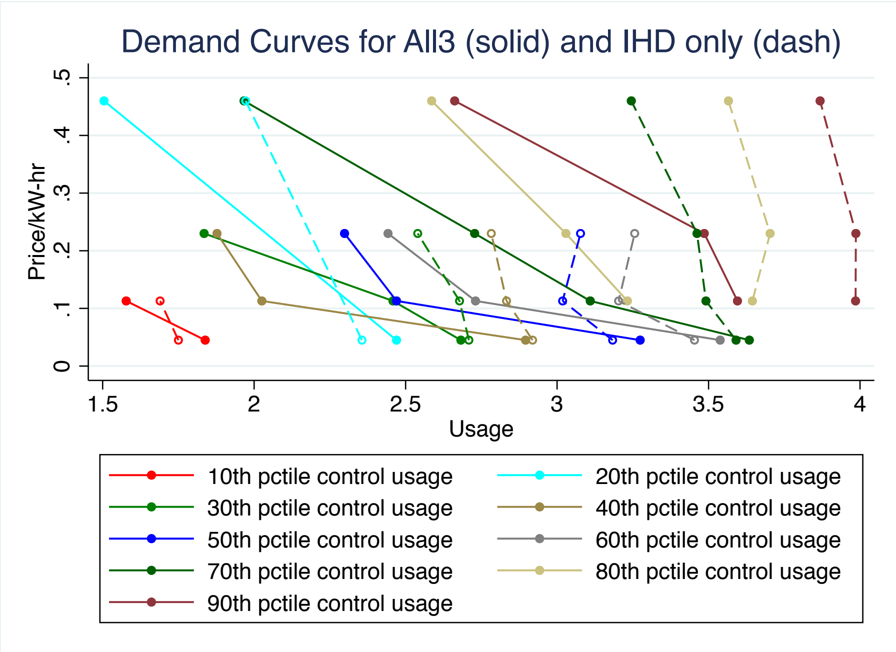


Figure 12: All3 vs. IHD Demand: All Deciles of Control Demand

Appendix F: Best Bill Evaluation

One possible confound to our inference of elasticities is that participants in the study were given a best bill guarantee. This means that customers who spend more under their treatment than they would have if the consumption was priced under the standard tiered structure are actually charged based on that tiered structure instead of the TOU or VPP variable rates. This itself could generate inelasticity if the variable rate structure will not apply.

We focus our discussion of the problem and analysis on customers under the VPP rate plan. The primary reason for this is that the focal measurement of elasticities in this paper is to compare PCT and non-PCT treated consumers based on the price variation arising under the VPP plan. The TOU plan does not provide dynamic price variation (i.e. it is a fixed plan over time with different prices for different times of days or days of week). Further, the VPP price plan has more frequent high peaks in prices (i.e. the 46c critical price is charged in non-critical times). The best bill criteria binds when the variable price realizations within a month are higher and consumers do not sufficiently respond to those prices to avoid their bills inflating above what they would have paid if the traditional tiered rate structure applied. To be precise, customers are not charged under the variable rate structure when the following condition holds:

$$R_1 > R_0$$
$$\sum_h P_h K_h > \sum_h p_c K_h$$

where R_1 indexes total revenue under treatment, R_0 is revenue if untreated and p_c is the control group price that follows the tiering structure where $p_c = 8.4c$ if $kWh < 1,400$ and $9.68c$ otherwise.

Across the three months in our data, the variable rate structure applied (i.e. $R_1 < R_0$) to more than 99.5 percent of bills in July and September. August had more days with the maximum 46c price; 43% of bills were set based on the control group pricing because $R_1 > R_0$.[15] If these customers in August anticipated this was going to be the case, they may have paid less attention to price notifications and/or altered settings on automated thermostats such that there is increased inelasticity to price. To evaluate this possibility, we rerun our analysis from the last column in Tables 6 and 7 excluding August and focusing only on August. The results for non-August resemble those in the paper. There is statistically significant but very modest elasticity of -0.035 for customers without either information or communication technology (i.e. the Portal is how they can look up prices). There is no added elasticity to having communications technology (i.e. the IHD), but once customers get an automation device elasticities jump up to -0.238 which is quite close to that reported in the main body of the paper in Table 7. During August, the baseline elasticity without communications or automation is no longer statistically significant and is slightly smaller than outside of August. The automation technology (PCT and All3) provides August elasticities almost identical to non-August. This suggests that if the best bill guarantee is influencing elasticity, it is only doing so to a very marginal degree for customers without automation technology.

Next, we replicate the non-parametric estimation of the demand curves excluding August. Figure 13 makes it clear that IHD and Portal demand curves are either perfectly inelastic or incredibly close to that throughout all price points tested in the data. Given these results for the non-August observations when best bill constraints did not bind, we believe our overall findings are not the result of the best bill guarantee that was offered.

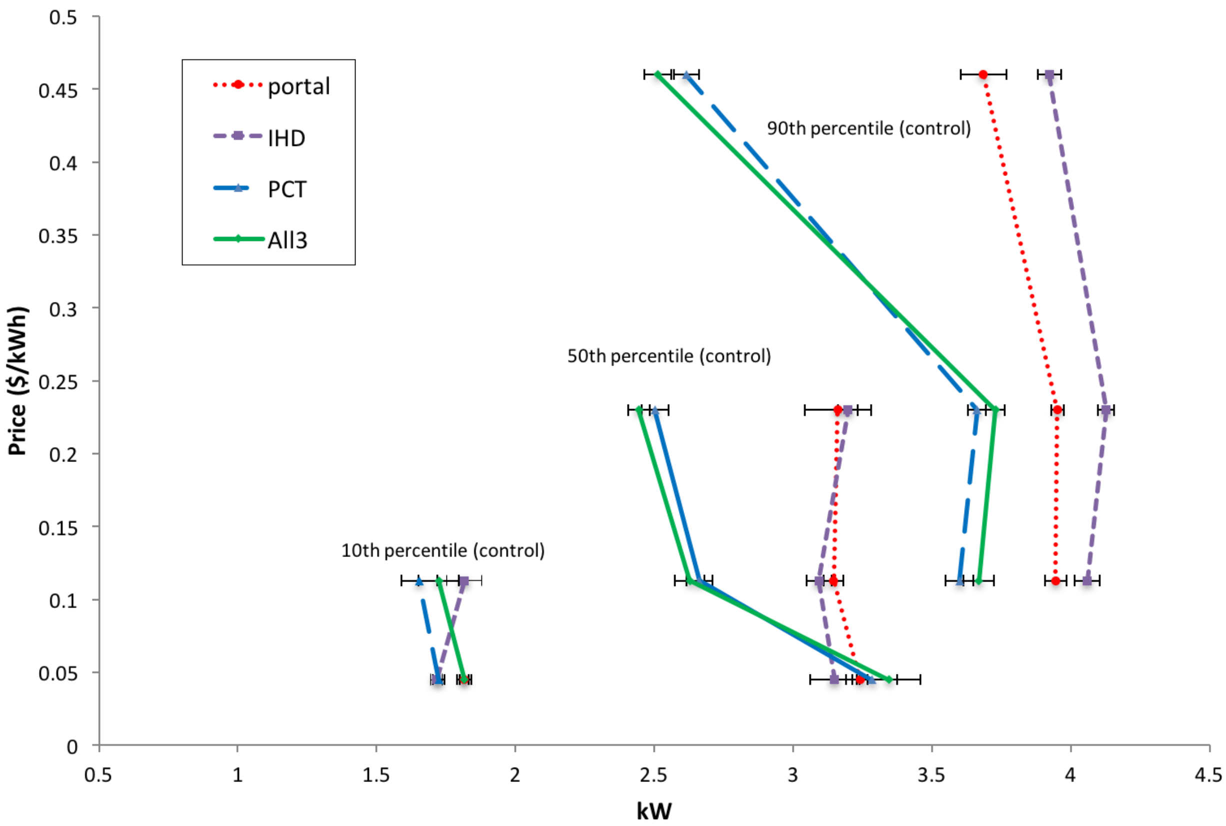
[15] Notably average demand from the control group was actually higher in July, but VPP prices in July never exceeded 23c.

Table 14: Control Function Regressions of Price Response: VPP Only

	Linear		Log-Log	
	Not August	**August Only**	**Not August**	**August Only**
Price	-0.484**	-0.161	-0.035**	-0.017
	(0.169)	(0.170)	(0.012)	(0.013)
Price*IHD	-0.170	-0.084	-0.016	-0.004
	(0.199)	(0.185)	(0.012)	(0.015)
Price*PCT	-2.726***	-2.096***	-0.238***	-0.174***
	(0.270)	(0.232)	(0.022)	(0.020)
Price*All3	-2.743***	-2.652***	-0.221***	-0.227***
	(0.289)	(0.264)	(0.022)	(0.021)
Fixed Effects	Meter	Meter	Meter	Meter
R-squared	0.625	0.700	0.616	0.668
N	289,809	104,547	289,809	104,547

Standard errors are clustered by meter ID and bootstrapped for control function specifications to accommodate error in the generated regressor.The latter two specifications substitute a control function interacted with technology treatment for the technology-hour-fixed effects. *** p=0.001, ** p=.01%, * p=.05.

Figure 13: All3 vs. IHD Demand: 10th, 50th and 90th Percentiles of Control Demand - No August Data



Appendix G: Non-Parametric Identification Argument

Our identification challenge is that we observe the joint distribution (Y, P) which is between two endogenous variables determined by a demand model, $Y = q_d(P, \eta)$, and a supply model $P = q_s(Y, \omega)$ where η and ω are the exogenous unobservables in the system underlying the distributions we observe in the data. Our object of interest is an elasticity based on the derivative $\partial Y / \partial P$ as generated through the demand model $q_d(\cdot)$. We cannot recover the causal relationship between Y and P to do this because P is correlated with η through the supply model $q_s(\cdot)$. To recover the causal relationship generated by $q_d(\cdot)$, we therefore need data providing a joint distribution of Y and P, conditional on η, i.e. we need the distribution $(Y, P|\eta)$.

The experiment in our data randomly held out some households to a control group that stayed on what is practically a fixed pricing plan we will characterize as $P = p_0$. For these households, we observe $Y_0 = q_d(P = p_0, \eta)$. From Y_0, we can therefore recover the marginal distribution (η). Now, when we inspect the households who were not held out from the endogenous price variation (i.e. a treated group), we observe $Y_1 = q_d(P, \eta)$. Combining the data from the treatment and control groups, we observe the joint distribution (Y_1, P, Y_0), which by inverting the function $q_d(\cdot)$ to recover η from Y_0 , we can obtain the joint distribution (Y_1, P, η). Together with the marginal distribution η, we can obtain the conditional distribution $(Y_1, P|\eta)$, which allows us to understand the derivative $\partial Y / \partial P$ as generated through $q_d(\cdot)$.

The relevance of ω in $q_s(\cdot)$ that, without it, there would be no variation in P conditional on η. In other words, the joint distribution $(Y_1, P|\eta)$, which is our identifying relationship essentially comes from the joint distribution $(Y_1, \omega|\eta)$. This has a similar identifying intuition of an instrumental variable (IV). Supply side shocks are exogenous to the demand system, but create variation in price that traces out the demand effect. In an IV approach, the demand relationship is estimated from a distribution of data (Y_1, Z), where Z creates variation in price that traces out the demand effect. Thus, we are actually better off than using an IV because we would only use the observable subset Z of ω.

For implementation purposes, P is not continuous and we do not try to formally transform the data into a conditional distribution $(Y_1, P|\eta)$. Rather, we condition on η, by conditioning on the levels of Y_0 we observe in the data. We then estimate the average demand at each price point observed for a given level of Y_0 .