Selection on Welfare Gains: Experimental Evidence from Electricity Plan Choice

Koichiro Ito, Takanori Ida, and Makoto Tanaka
JANUARY 2021

EPIC

ENERGY POLICY INSTITUTE
AT THE UNIVERSITY OF CHICAGO

An Affiliated Center of

Becker
Friedman
Institute

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

SELECTION ON WELFARE GAINS:
EXPERIMENTAL EVIDENCE FROM ELECTRICITY PLAN CHOICE

Koichiro Ito
Takanori Ida
Makoto Tanaka

January 2021

Selection on Welfare Gains: Experimental Evidence from Electricity Plan Choice
Koichiro Ito, Takanori Ida, and Makoto Tanaka
January 2021
JEL No. L94,Q41

ABSTRACT

We study a problem in which policymakers need to screen self-selected individuals by unobserved heterogeneity in social welfare gains from a policy intervention. In our framework, the marginal treatment effects and marginal treatment responses arise as key statistics to characterize social welfare. We apply this framework to a randomized field experiment on electricity plan choice. Consumers were offered socially efficient dynamic pricing with randomly assigned take-up incentives. We find that price-elastic consumers—who generate larger welfare gains—are more likely to self-select. Our counterfactual simulations quantify the optimal take-up incentives that exploit observed and unobserved heterogeneity in selection and welfare gains.

Koichiro Ito
Harris School of Public Policy
University of Chicago
1155 East 60th St
Chicago, IL 60637
and NBER
ito@uchicago.edu

Takanori Ida
Graduate School of Economics
Kyoto University
Yoshida, Sakyo
Kyoto 606-8501
Japan
ida@econ.kyoto-u.ac.jp

Makoto Tanaka
National Graduate Institute for Policy Studies
7-22-1 Roppongi, Minato- ku,
Tokyo 106-8677
Japan
mtanaka@grips.ac.jp

1 Introduction

Selection is a key phenomenon in economic policies because voluntary take-up is a ubiquitous feature of policy design. For example, social safety net programs in the United States, such as food stamps and disability insurance programs, require eligible individuals to voluntarily apply for benefits (Finkelstein and Notowidigdo, 2019; Deshpande and Li, 2019). A similar design is widely used in other policy domains, including worker training programs (LaLonde, 1986), childcare programs (Cornelissen et al., 2018), Medicare payment programs (Einav et al., 2020), energy efficiency rebates (Borenstein and Davis, 2016; Allcott and Greenstone, 2017), and insurance and utility service plans (Einav et al., 2013; Hortaçsu, Madanizadeh and Puller, 2017; Fowlie et al., 2017).

In the presence of selection, a policy's social welfare gains depend on what we call *selection on welfare gains*. Selection is governed by each individual's private gain from the policy. This private gain itself is an important outcome, but what matters to policymakers is the social welfare gain that can be generated by each individual's take-up and behavioral response to the policy intervention. If this welfare gain is heterogeneous among individuals, the connection between private selection and social welfare gains plays an important role in policy design.

We formalize this idea by developing a framework that connects heterogeneity in private gains and heterogeneity in social welfare gains. Our framework builds on the principles of welfare economics and the generalized Roy model developed by Eisenhauer, Heckman and Vytlacil (2015). An individual selects into treatment based on a selection equation, which is a function of observables and unobservables. The Roy model allows us to characterize heterogeneity in potential outcomes using the marginal treatment effect (MTE) and the marginal treatment responses (MTRs), as shown by Mogstad, Santos and Torgovitsky (2018). We combine this framework with the principles of welfare economics to show that heterogeneity in social welfare gains can be expressed by a function of the MTE, the MTR, and the selection equation. With this social welfare function in hand, we can examine a policy's social welfare gain in the presence of selection.

An advantage of this framework is that it clarifies the parameters that need to be estimated through empirical analysis. With this insight, we designed a randomized controlled trial (RCT) in the field to study electricity plan choice, where the policy goal was to generate social welfare gains from the adoption of dynamic electricity pricing. Similar to nearly all households in the United

States and the rest of the world, consumers in our experiment had a non-dynamic electricity price (i.e., the price did not vary by hour). This commonly used pricing is socially inefficient because the marginal cost of electricity tends to vary substantially by hour (Joskow and Wolfram, 2012). To address this inefficiency, we offered customers a dynamic pricing plan. In addition, we offered them a randomly assigned financial incentive for take-up. We use this random variation in take-up incentive as an instrument to estimate the MTE and MTRs based on a method developed by Brinch, Mogstad and Wiswall (2017).

We begin by estimating the selection equation. Before the field experiment, we collected demographic information and solicited household-level risk preferences based on the method developed by Callen et al. (2014). We also collected historical hourly usage data at the household level. These data allowed us to calculate each consumer's expected saving from dynamic pricing, that is, the financial saving from dynamic pricing under the assumption of no behavioral response to price changes. We find that selection is strongly related to the expected savings from dynamic pricing. This provides empirical evidence of "selection on the level" that has been documented in the healthcare markets (Einav et al., 2013). We also find that selection is positively associated with the take-up incentive (Z) and years of schooling, but negatively related to risk aversion, certainty premium, and employment.

The propensity score that is obtained from the selection equation plays a central role in estimating the MTE and MTRs. We estimate the causal effect of price changes on consumption (i.e., the price elasticity of demand) for peak and off-peak hours. The advantage of the MTE is that it allows us to estimate the relationship between the average treatment effects (ATE) and selection. We use the MTE to test "selection on the slope" (Einav et al., 2013), that is, whether more-elastic consumers are more likely to adopt dynamic pricing. We find strong evidence of selection on the slope for demand in peak hours. The estimated MTE function also indicates that price elasticity diminishes to zero as we approach consumers less likely to adopt dynamic pricing. Because the social welfare gain from dynamic pricing is directly tied to the slope of demand, our result implies that the marginal social welfare gain is likely to diminish as we approach consumers less likely to adopt.

We investigate this question by plugging our empirical estimates of the selection equation, MTE, and MTRs into the social welfare function developed in our framework. This approach allows us

to conduct three welfare analyses. First, we quantify the social welfare gains from the policies implemented in our field experiment. We find that the take-up incentive provided in our experiment increased social welfare.

Second, we conduct counterfactual policy simulations to find the optimal take-up incentive. We find that the welfare gain is increasing in the take-up incentive (Z) up to a certain level, but the relationship is concave for two reasons. First, the estimated MTE function suggests that the treatment effect (i.e., the behavioral responses to the price change) diminishes as we increase the take-up incentive to induce consumers who would not adopt dynamic pricing in the absence of a high take-up incentive. Second, these consumers, who need a larger take-up incentive, are those who have larger unobserved disutility from adopting dynamic pricing. For these reasons, the marginal welfare gain from increasing the take-up incentive diminishes to zero.

We show that this is empirically the case and estimate the optimal level of the take-up incentive for two scenarios. The first optimal incentive is the optimal uniform take-up incentive ($Z = z^*$), where policymakers cannot differentiate the incentive by observables. The second one is the optimal differentiated take-up incentives ($Z = z^*(x)$), under which policymakers can differentiate the incentive by observables. Recall that our welfare function is a function of the MTE, MTR, and propensity score from the selection equation. Therefore, the differentiated take-up incentive can exploit variation in observables in both the selection equation and the treatment effects.

We compare the welfare gains from four policies: i) $Z = 0$, ii) $Z = 60$, iii) $Z = z^*$, and iv) $Z = z^*(x)$. The welfare gains from the first two policies are calculated directly from the variation in the experiment. Those from the last two policies are based on counterfactual policy simulations. We find that $Z = 60$ increases welfare relative to $Z = 0$ and that there are additional welfare gains by setting Z at the optimal level. This welfare gain can be further enhanced by exploiting variation in observables and using the optimal differentiated take-up incentives ($Z = z^*(x)$).

Finally, we use our framework to discuss implications for a mandatory take-up policy. The vast majority of countries rely on voluntary take-up policies for the adoption of dynamic electricity pricing for residential customers because a mandate is politically infeasible. However, it is still useful to discuss what conditions could make a mandate more welfare-enhancing than the counterfactual policies discussed above. A key unknown effect of a mandate is how it changes a consumer's disutility from adopting dynamic pricing. Because this effect is unknown from our experiment

and previous studies, we examine two possibilities. First, we show that if a mandate does not change such disutility, our counterfactual simulation results imply that it is inferior to the optimal take-up incentive policies because additional take-up beyond the optimal level would reduce welfare. Second, if a mandate can reduce disutility for adopting dynamic pricing, we can use our framework to calculate how much such a benefit has to be to make a mandate policy more welfare-enhancing than other policies. Our calculation indicates that a mandate can be superior to other counterfactual policies if its net benefit, including the political cost of implementing it, is more than \$26 per customer per year.

Related literature and our contributions—First, our framework builds on the literature on the MTE and the policy-relevant treatment effects (Heckman and Vytlacil, 2001, 2005; Heckman, 2010). The underlying concept in this literature is motivated by Marschak (1953), which states that for many policy analyses, it may not be necessary to identify fully specified structural models. Instead, researchers may be able to address policy questions by estimating a set of relevant structural parameters derived from economic models. Recently, Mogstad, Santos and Torgovitsky (2018) extends this concept by introducing a so-called *target parameter*. A target parameter is any function of MTRs, and therefore, encompasses the MTE. We show that the social welfare function in our study emerges as a target parameter. Our counterfactual policy simulations highlight that this tool can be useful for analyzing not only the private benefits from a policy intervention but also its implications for social welfare.[1]

Second, our study is also related to sufficient-statistics approaches (Chetty, 2009; Kleven, 2020), which is another strand of the literature influenced by Marschak (1953). In this stream of the literature, researchers use economic theory to derive a social welfare function with parameters that can be estimated from experimental or quasi-experimental variation in data. Typical sufficient-statistics approaches do not consider selection. Without selection, the *average* treatment effect (ATE) is sufficient to characterize a policy's welfare gain, even if the treatment effects are heterogenous. However, many policies may have both selection and heterogeneous treatment effects. Our analysis shows that these two elements can be incorporated into the framework of sufficient-statistics approaches by estimating the selection equation, the MTE, and MTRs. While Chetty (2009) recognizes this key

---

[1]The spirit of our welfare analysis based on the MTE is similar to Kline and Walters (2016), which evaluates policy counterfactuals of the Head Start program using the MTE.

connection between sufficient-statistics approaches and the MTE, our study is the first to develop an empirical framework that makes this connection.

Third, our study is closely related to the recent literature on policy take-up and targeting (Finkelstein and Notowidigdo, 2019; Deshpande and Li, 2019). These studies show that optimal targeting requires knowledge of how selection in take-up is related to individual characteristics. Our study contributes to this body of the literature by showing that policymakers need additional key information—how selection in take-up is related to heterogeneity in the treatment effects and the resulting social welfare gains. To investigate this point, we need to estimate both selection and the MTE from a policy intervention. Our experimental design allows us to conduct such an analysis, and we show that the connection between selection and heterogeneity in social welfare gains plays a key role in targeting.[2]

Another closely related stream of the literature is research on selection outside the electricity industry, such as in healthcare sectors. Einav et al. (2013) is the first study to decompose selection in the context of health insurance into two key elements: selection on the *level* and selection on the *slope*. Since then, both types of selection have been recognized as key phenomena in the study of healthcare markets (Einav et al., 2020). Our study highlights that this concept also plays a key role in energy markets. A notable distinction is the implication of selection on the slope. In healthcare markets, selection on the slope is generally considered to be negative for social welfare because it exaggerates moral hazard. However, such selection can be welfare-enhancing in corrective policies such as the adoption of dynamic electricity pricing. This is because in this situation, the social planner benefits more from consumers with more-elastic demand.

Finally, our analysis provides important implications for ongoing discussions in energy policy in many countries. Households in the United States and many countries now have smart meters at home that record hourly consumption and, therefore, make dynamic pricing technologically feasible (Joskow, 2012). In fact, recent studies show that there is substantial efficiency gains if residential consumers adopt dynamic pricing (Wolak, 2011; Ito, Ida and Tanaka, 2018). However, the mandate of dynamic pricing is politically infeasible for most countries, and therefore, policymakers need to design a policy with voluntary take-up. Our result suggests that it can be possible to improve

---

[2]Conceptually, our targeting strategy is similar to that of Kitagawa and Tetenov (2018), in which targeting is based on social welfare gains from a policy intervention instead of individual characteristics per se. Kitagawa and Tetenov (2018) focus on mandatory policy assignments without selection.

welfare by incentivizing consumers to adopt dynamic pricing. However, there is also likely to be a diminishing return to such policies. As we show in our welfare analysis, consumers who are less likely to adopt dynamic pricing are likely to be relatively price-inelastic consumers, who may not generate large social welfare gains from the adoption of dynamic pricing.

2 Background

In this section, we describe fundamental inefficiency in electricity markets, why dynamic pricing can mitigate this problem, and why selection is the key factor to policy design.

2.1 Fundamental Inefficiency in Electricity Markets

Inefficient retail pricing has long been a central issue in electricity markets (Joskow, 2012). In the United States, nearly all households pay electricity prices that do not reflect the marginal cost of electricity (FERC, 2011). The cost of electricity is time-varying because of substantial fluctuations in demand and a lack of storability (Borenstein, 2002). Therefore, a socially efficient electricity tariff would be dynamic pricing that reflects the time-varying marginal cost.

However, typical residential electricity prices are time-invariant. The wedge between the marginal cost and marginal price creates a textbook example of deadweight loss. Compared with the efficient level, households consume too much when marginal costs are higher than retail rates—typically in peak demand hours—and too little when marginal costs are lower than retail rates—typically in off-peak hours. This phenomenon is not unique to the United States and is ubiquitous in most countries, except for a few nations in Europe.

[Figure 1 about here]

Figure 1 illustrates this inefficiency for two hypothetical demand curves. Consider a change from flat pricing to dynamic pricing. The social welfare gain from dynamic pricing is Harberger's triangle in the figure. Importantly, the size of the welfare gain depends on the slope of demand. When a more-elastic consumer adopts dynamic pricing, the welfare gain is $A + B$ in peak hours and $C + D$ in off-peak hours. The welfare gain from a less-elastic consumer is B in peak hours and C in off-peak hours. Thus, the more-elastic consumer would produce a larger welfare gain from adopting dynamic pricing.

6

2.2 Who Selects into Dynamic Pricing?

The heterogeneity in welfare gains presented in Figure 1 plays a key role in policy design because residential dynamic electricity pricing is usually offered through voluntary take-up. With voluntary take-up, a key question is what types of consumers are likely to select into dynamic pricing. Before we provide a formal model in the next section, we describe potential selection mechanisms.

[Figure 2 about here]

The first potential selection mechanism is *selection on the slope* (Einav et al., 2013). Figure 2 describes this theoretical prediction, showing consumers' private gains from adopting dynamic pricing with two hypothetical demand curves. When the *more-elastic* consumer adopts dynamic pricing, the private gain (the change in consumer surplus in this figure) is $-E$ in peak hours and $G + H$ in off-peak hours. The private gain for the *less-elastic* consumer is $-(E + F)$ in peak hours and G in off-peak hours. Thus, the more-elastic consumer would have a larger private gain from selecting into dynamic pricing in terms of the consumer surplus.

The second selection mechanism is *selection on the level* (Einav et al., 2013). Consider two consumers who have the same price elasticity, and therefore, have the same selection on the slope. One consumer uses more electricity in off-peak hours, whereas the other consumer uses more electricity in peak hours. The private gain from dynamic pricing is larger for the consumer who uses more in off-peak hours because dynamic pricing increases the peak price and lowers the off-peak price. Such consumers are called the *structural winners* of dynamic pricing. In Figure 2, selection on the level is related to the *intercepts* of the demand curves, whereas selection on the slope is related to the slopes of the demand curves. As we see in Section 4, a typical dynamic electricity tariff creates a wide distribution of structural winners and losers because of the variation in electricity usage in peak and off-peak hours. Therefore, selection on the level can be another key selection mechanism in our context.

The third mechanism is through the *subjective cost of treatment* (Eisenhauer, Heckman and Vytlacil, 2015). Consider two consumers with identical electricity demand, including their slope and intercept. Then, they have the same selection on the slope and level. However, their selection can still differ for at least two reasons. First, they may have differences in inertia or switching costs to change their electricity plans. Second, they may have differences in unobserved plan preferences

that cannot be explained by observable plan characteristics such as electricity prices. In the terminology of Eisenhauer, Heckman and Vytlacil (2015), both elements are part of the subjective cost of treatment in the selection equation, as described in the next section. Our framework below incorporates this subjective cost of treatment.

3 Conceptual Framework

In this section, we formalize the selection mechanisms and social planner's problem discussed in Section 2. Our framework has two goals. First, we use welfare economics and the generalized Roy model to connect heterogeneity in private gains and heterogeneity in social welfare gains from a policy. Second, we show that, in our context, the social welfare function in the presence of selection can be written as a *target parameter*, as defined by (Mogstad, Santos and Torgovitsky, 2018). This framework clarifies the parameters that have to be empirically estimated to conduct the welfare analysis.

Figure 3 provides an overview of our framework. The social planner can provide a financial take-up incentive (Z) for consumers to adopt dynamic pricing. Consumers self-select based on the selection equation in Section 3.1. In Section 3.2, we show that three parameters—the MTE, MTRs, and the propensity score from the selection equation—arise as key statistics to characterize social welfare in the presence of selection. We provide visual summary of the key results of this framework in Section 3.3 and explain how this framework relates to our empirical analysis in Section 3.4.

[Figure 3 about here]

3.1 Consumer's Problem

Consider an electricity consumer who has a default electricity price plan ($j = 0$) and an option to adopt a dynamic price plan ($j = 1$). A binary variable $D \in \{0, 1\}$ equals one if the consumer chooses dynamic pricing. We use Y_t to denote the consumer's hourly electricity consumption in hour $t = 1, ..., T$. The relationship between the observed outcome (Y_t) and potential outcomes ($Y_{t,j}$) for $j \in \{0, 1\}$ is

$$Y_t = DY_{t,1} + (1 - D)Y_{t,0}. \tag{1}$$

The consumer obtains indirect utility S_j from choice j and selects into dynamic pricing if $S_1 - S_0 > 0$. Following Heckman and Vytlacil (1999, 2005) and Mogstad, Santos and Torgovitsky (2018), we assume that the net surplus $S_1 - S_0$ is weakly separable between observables and unobservables:

$$S_1 - S_0 = \nu(X, Z) - V. \tag{2}$$

$\nu(X, Z)$ is the observable part of the selection equation with a flexible function $\nu(.)$, observables X, and an instrument Z that affect selection. V is the unobserved disutility for selecting $j = 1$. We make the following assumptions.

Assumption 1. $D = 1$ if $S_1 - S_0 > 0$.

Assumption 2. $(Y_{t,0}, Y_{t,1}, V) \perp\!\!\!\perp Z|X$, where $\perp\!\!\!\perp$ denotes conditional independence.

Assumption 3. V is continuously distributed, conditional on X.

Assumption 1 means that consumers self-select based on the selection equation. Assumption 2 implies that conditional on X, instrument Z is independent of the potential outcomes and unobservables in the selection equation. Assumption 3 requires the distribution of V to be continuous, but it does not impose a particular distributional assumption. We use $F_V(.)$ to denote the cumulative distribution function of V. Then equation (2) implies that the propensity score is $P(X, Z) \equiv \Pr[D = 1|X, Z] = F_V(\nu(X, Z))$.

The key variable we define is $U \equiv F_V(V) \in [0, 1]$. Because V is the unobserved *disutility* for choosing dynamic pricing, U tells us the quantiles of consumer types in terms of their unobserved preference. Consumers with lower U are more likely to adopt dynamic pricing for unobservable reasons. Consumers with higher U are less likely to adopt dynamic pricing for unobservable reasons.

With this setup, we can define the **Marginal Treatment Effect (MTE)** following Heckman and Vytlacil (1999, 2005):

$$Y_t^{MTE}(x, u) = E[Y_{t,1} - Y_{t,0}|X = x, U = u]. \tag{3}$$

One interpretation of the MTE is that it is a function of the ATE in observables $X = x$ and unobservables $U = u$. The MTE is useful for our research question because it allows a direct test of selection on the slope (Einav et al., 2013) in terms both of the observables and unobservables.

First, consider consumer types whose observables X make them more likely to adopt dynamic pricing through the selection equation in (2). The link between the MTE and X allows us to test if these observable types have larger or smaller behavioral responses to treatment (i.e., changes in electricity consumption in response to changes in electricity price) than others.

Second, consider consumer types whose unobservables U make them more likely to self-select, conditional on observables X. That is, their values of U—the unobserved disutility for adopting dynamic pricing—are lower than others. The link between the MTE and U allows us to test if these unobservable types have larger or smaller behavioral responses to treatment than others. The advantage of this test is that it does not have to impose distributional assumptions on the unobservables, except for Assumptions 1, 2, and 3.

Other useful estimands for our study are the **Marginal Treatment Responses (MTRs)** introduced by Mogstad, Santos and Torgovitsky (2018):

$$m_{t,0}(x,u) = E[Y_{t,0}|X = x, U = u]$$
$$m_{t,1}(x,u) = E[Y_{t,1}|X = x, U = u]. \tag{4}$$

The MTRs are simply two components of the MTE. There are at least two advantages of working with the MTRs. First, $m_{t,0}(x,u)$ provides the average untreated outcome for consumer types $U = u$ and $X = x$. Therefore, $m_{t,0}(x,u)$ provides important information on selection in terms of the untreated potential outcome (Einav, Finkelstein and Schrimpf, 2010). Second, this allows us to consider policy-relevant target parameters that depend on $m_{t,0}$ and $m_{t,1}$ asymmetrically. This is particularly important in our study because we show below that some of our policy-relevant target parameters are asymmetric functions of $m_{t,0}$ and $m_{t,1}$.

3.2 Social Planner's Problem

We now consider the social planner's problem and its key connection to the consumer's problem. As shown in Figure 1, the adoption of dynamic pricing generates a social welfare gain equal to the deadweight loss of time-invariant pricing. This welfare gain can be expressed by potential outcomes

for social welfare, consumer surplus, and producer surplus as follows:

$$W_1 - W_0 = (S_1 - S_0) + (PS_1 - PS_0), \tag{5}$$

where (W_j, S_j, PS_j) are potential outcomes for social welfare (W), consumer's indirect utility (S), and producer surplus (PS) for the choice of price plan $j \in \{0, 1\}$.

This allows us to define the MTE of the social welfare gain for consumer types $U = u$ and $X = x$:

$$
\begin{aligned}
W^{MTE}(x, z, u) &= E[W_1 - W_0 | X = x, Z = z, U = u] \\
&= S^{MTE}(x, z, u) + PS^{MTE}(x, z, u), \tag{6}
\end{aligned}
$$

where X is observables, Z is a take-up incentive, and U is unobserved disutility for adopting dynamic pricing. Take-up incentive Z is a transfer between consumers and producers; therefore, although it affects a consumer's selection decision, the transfer will be canceled out in the calculation of social welfare. We assume that the indirect utility function is quasi-linear. Then, the selection equation (2) implies that $S^{MTE}(x, z, u) = E[S_1 - S_0 | X = x, Z = z, U = u] = \nu(x) + z - F_V^{-1}(u)$. Using the definition of producer surplus, W^{MTE} can be written as[3]

$$
\begin{aligned}
W^{MTE}(x, z, u) &= E[W_1 - W_0 | X = x, Z = z, U = u] \\
&= \nu(x) - F_V^{-1}(u) + \sum_t \left[(\tau_{t,1} - c_t) \cdot m_{t,1}(x, u) - (\tau_{t,0} - c_t) \cdot m_{t,0}(x, u) \right], \tag{7}
\end{aligned}
$$

where $\tau_{t,j}$ and $c_{t,j}$ are the marginal price and marginal cost of electricity at time t for price plan j, respectively.[4] This key result implies that the MTE of the social welfare gain $(W^{MTE}(x, z, u))$ can be expressed by a function of the parameters in the selection equation: $\nu(x) - F_V^{-1}(u)$, and the MTRs: $m_{t,0}(x, u)$ and $m_{t,1}(x, u)$.

The social planner's primary objective is to maximize the welfare gain per capita. In the absence

---

[3]Note that $PS_1 - PS_0 = \sum_t \left[(\tau_{t,1} - c_t) \cdot Y_{t,1} - (\tau_{t,0} - c_t) \cdot Y_{t,0} \right] - z$. Therefore, $PS^{MTE}(x, z, u) \equiv E[PS_1 - PS_0 | X = x, Z = z, U = u] = \sum_t \left[(\tau_{t,1} - c_t) \cdot m_{t,1}(x, u) - (\tau_{t,0} - c_t) \cdot m_{t,0}(x, u) \right] - z$.

[4]In this model, we consider a constant marginal cost c_t given time t. This is consistent with our empirical context because the change in consumption from customers in our experiment was unlikely to change the very short-run marginal cost of electricity. For different applications, the model can be easily modified to incorporate the case in which the marginal cost at time t is a step function of electricity demand at time t.

of selection, the planner's objective function is simply the policy's ATE on the welfare gain. However, in the presence of selection, the objective function becomes the intention-to-treat (ITT) on the welfare gain:

$$W^{ITT}(x,z) \equiv \int_0^1 W^{MTE}(x,z,u)du = \int_0^{P(x,z)} W^{MTE}(x,z,u)du + \int_{P(x,z)}^1 W^{MTE}(x,z,u)du$$

$$= \int_0^{P(x,z)} W^{MTE}(x,z,u)du, \tag{8}$$

where $P(x,z)$ is the propensity score. The last equality comes from the fact that $D = 0$ for $P(x,z) \le u \le 1$ so that $W^{MTE}(x,z,u) = 0$ for these consumers.

Two additional statistics are policy-relevant. The first statistic is the average treatment effect on the treated (ATET) on the welfare gain: $W^{ATET}(x,z) \equiv W^{ITT}(x,z)/P(x,z)$. This calculates the policy's ATE on the welfare gain for those who self-select given the value of Z.

Another policy-relevant statistic is the marginal welfare gain with respect to the policy instrument $(dW^{ITT}(x,z)/dz)$, which measures how much marginal welfare gain can be obtained by increasing the value of policy instrument Z. This parameter is often the primary interest in the sufficient statistics literature. Chetty (2009) and Kleven (2020) show that even if a policy's treatment effect is heterogenous among individuals, the ATE is often sufficient to describe the policy's marginal welfare gain in the absence of selection. The following equation shows that in the presence of selection, the marginal welfare gain becomes a function of $W^{MTE}(x,z,u)$ and propensity score:

$$\frac{dW^{ITT}(x,z)}{dz} = \frac{d}{dz}\int_0^{P(x,z)} W^{MTE}(x,z,u)du$$

$$= \underbrace{\frac{dP(x,z)}{dz}}_{\substack{\text{Size of marginal} \\ \text{consumers}}} \cdot \underbrace{W^{MTE}(x,z,P(x,z))}_{\substack{\text{Welfare gain from} \\ \text{marginal consumers}}}. \tag{9}$$

The second equality comes from the Leibniz integral rule. Equation (9) can be seen as a generalization of the sufficient statistics approach, in which selection is incorporated. This equation implies that in the presence of selection, the MTE of the welfare gain $(W^{MTE}(x,z,u))$ and the propensity score $(P(x,z))$ from the selection equation become key statistics to characterize the marginal welfare gain with respect to policy instrument Z.

Equation (8) also implies that the socially optimal Z can be obtained by estimating propensity

score $P(x, z)$ and the MTE of the welfare gain $W^{MTE}(x, z, u)$. When the social planner can differentiate Z based on observables X, the planner can find the optimal differentiated take-up incentive $z^*(x)$ by

$$z^*(x) = \underset{z(x)}{\operatorname{argmax}} \int_0^{P(x,z)} W^{MTE}(x, z, u) du. \tag{10}$$

When the planner cannot differentiate Z by X, the planner can find the optimal uniform take-up incentive z^* by

$$z^* = \underset{z}{\operatorname{argmax}} \int_X \int_0^{P(x,z)} W^{MTE}(x, z, u) du \ dF_X \tag{11}$$

where F_X is the distribution of X.

3.3 Link Between Selection and Social Welfare Gains

Figure 4 summarizes the key results of our framework by illustrating the connection between a consumer's selection problem and the social planner's problem. The horizontal axis is $U \equiv F_V(V) \in [0, 1]$, which is the quantiles of unobserved consumer heterogeneity in terms of their disutility for adopting dynamic pricing. Conditional on observables X, consumers with lower values of U (i.e., the left-hand side of the horizontal axis) are more likely to adopt dynamic pricing and vice versa.

[Figure 4 about here]

Panel A illustrates a hypothetical example of the selection equation, $S_1 - S_0 = \nu(x, z) - V$, for given $X = x$ and $Z = z$. The selection equation is downward sloping in this figure because U is the quantiles of V. That is, given $X = x$ and $Z = z$, consumers with lower values of U have lower values of V and thus they are more likely to adopt dynamic pricing. More precisely, given take-up incentive $Z = z$, consumers with $0 \leq U < P(x, z)$ adopt dynamic pricing because $S_1 - S_0 > 0$. Consumers with $U \geq P(x, z)$ do not adopt because $S_1 - S_0 \leq 0$.

Panel B presents a hypothetical function of $W^{MTE}(x, z, u)$ as a function of U. The shape of $W^{MTE}(x, z, u)$ in this figure is hypothetical, and the empirically estimated function is flexible be upward-sloping, downward-sloping, or non-monotonic in U. The welfare gain per capita $W^{ITT}(x, z)$ is the integral of $W^{MTE}(x, z, u)$ in $0 \leq U < P(x, z)$. Therefore, $W^{ITT}(x, z)$ equals K in the figure.

In this example, take-up incentive $Z = z$ is not at the optimal level. The planner can set a higher incentive $Z = z^*$ based on equation (10). This optimal incentive moves the selection equation to the right $(S_1 - S_0 = \nu(x, z^*) - V)$. Then, consumers with $0 \leq U < P(x, z)$ adopt dynamic pricing, and $W^{ITT}(x, z)$ can be increased to $K + L$.

That is, if the social planner has knowledge of the selection equation and $W^{MTE}(x, z, u)$, it is possible to find the optimal level of Z by connecting the consumer's problem (Panel A) and the social planner's problem (Panel B). This visual interpretation is a summary of the key results in equations (9) to (11).

3.4 What Parameters Need to be Estimated?

Equations (9) to (11) clarify the parameters that need to be estimated in our empirical analysis to conduct the welfare analysis. First, we need to estimate the selection equation (2) to estimate $\nu(x) - F_V^{-1}(u)$, which also provides an estimate of propensity score $P(x, z)$. Second, we need to estimate the two MTRs $(m_{t,1}(x, u)$ and $m_{t,0}(x, u))$, which provide an estimate of the MTE. Using these estimates, we can construct an empirical estimate of $W^{ITT}(x, z)$ in equation (8), calculate $dW^{ITT}(x, z)/dz$, and find the optimal level of the take-up incentive.

With this insight, we designed our field experiment to generate the variation necessary to estimate these parameters. In particular, to estimate the MTE and MTRs, we need an instrumental variable that affects selection decisions but not the potential outcomes of electricity usage $(Y_{t,0}, Y_{t,1})$. In our experiment, we created an instrument by randomly assigning a financial take-up incentive. We explain the details of experiential design in the next section.

4 Experimental Design and Data

We conducted our field experiment in the city of Yokohama, Japan with the Japanese Ministry of Economy, Trade and Industry, the city of Yokohama, Tokyo Electric Power Company, Toshiba Corporation, and Panasonic Corporation. We collected household-level hourly electricity usage data in 2013 to calculate each customer's expected saving from adopting dynamic pricing and run the experiment in 2014 and 2015.

Although our experiment was an RCT, participating households were not a random sample of

the population, as was the case in previous experiments on electricity pricing (Wolak, 2006, 2011; Faruqui and Sergici, 2011). The city of Yokohama used online and offline public advertisements to recruit participating households. To recruit as broad ranges of households as possible, the city provided free installations of an advanced electricity meter and in-home display as well as a participation reward of JPY 20,000 ($\approx$ \$200) for all participating households. The city was able to recruit 3,293 households. We excluded students, customers who had self-generation devices, and those without access to the Internet. This process left us with 2,153 households. We used 1,183 customers for another field experiment (Ito, Ida and Tanaka, 2017), and used 970 customers for the current study. For these customers, we collected hourly electricity usage data at the customer level as the primary outcome variable for this study.

This sample selection is important when considering external validity. For this reason, we compare observable characteristics between households in our experiment and a random sample of households in the city in Table A.2. We do not find statistically significant differences in the observables between these two groups, including electricity usage and demographics. However, it is still important to recognize that our experimental sample can differ from the population in their unobservables.

4.1 Electricity Plan Choice

All the customers in our experiment were eligible to switch from their default electricity tariff to a dynamic pricing tariff. As shown in Figure 5, under the default flat pricing, the price was 26 cents per kWh regardless of the time of use. Under the dynamic pricing plan, the price in off-peak hours was reduced to 21 cents per kWh, and the price in peak hours was 45 cents per kWh on regular days and 100 cents per kWh on critical peak days.

[Figure 5 about here]

Peak hours were between 1 pm and 4 pm in summer and between 5 pm and 8 pm in winter corresponding to the system peak hours in the Japanese electricity system. Peak-hour prices were mostly 45 cents per kWh except for critical peak days, defined as those days for which the previous day's weather forecast predicted particularly high system demand relative to system supply. All customers received day-ahead and same-day notices about their prices via the in-home display and

15

text notifications.[5]

To generate the variation to estimate the MTE and MTRs, we randomly assigned a financial take-up incentive. Consumers in the treatment group were notified that they were going to receive an upfront incentive of JPY 6,000 ($\approx$ \$60) upon the take-up of dynamic pricing. As summarized in Table 1, the two groups differed only by the receipt of this take-up incentive.

[Table 1 about here]

4.2 Data and Descriptive Statistics

Using advanced electricity meters often called "smart meters," we collected hourly electricity usage data at the household level in the pre-experimental period in 2013 and experimental period in 2014 and 2015. The primary intervention (i.e., dynamic pricing) started on July 8, 2014 and continued until September 30 in the summer. The winter intervention started on December 1 and ended on January 31, 2015. Therefore, the sample size of hourly electricity usage data is 3,422,160 in the intervention period.[6]

In addition, we collected household demographics and elicited risk preferences using the method developed by Callen et al. (2014). The elicitation was based on two series of questionnaires (Table A.1). In the first task, a customer made a series of choices between a relatively safe option (option A) and a relatively risky option (option B). Where the customer switched from preferring option A to option B was used to determine the customer's risk aversion. In the second task, we added uncertainty into option A, but kept it less risky than option B to obtain another risk aversion parameter. The advantage of this method is that we can obtain two risk preference parameters, from which we can calculate the certainty premium as well as risk aversion.

[Table 2 about here]

Table 2 presents the summary statistics of the demographics and pre-experiment consumption data by treatment group. A comparison across the groups indicates statistical balance in the observables.

A key variable obtained from the pre-experimental data was the expected saving from dynamic pricing. We used each household's hourly usage data in the pre-experimental period to calculate a

---

[5]There were fourteen critical peak days in each of the summer and winter.

[6]This is because we have 970 customers × 147 days × 24 hours = 3,422,160. In our empirical analysis, we cluster the standard errors at the customer level to adjust for serial correlation.

counterfactual payment as if the customer was on dynamic pricing rather than default flat pricing. Using this value, we calculated each customer's expected annual saving from dynamic pricing based on the customer's past usage data. When we notified customers of their expected savings, we explicitly told them that this value was an *expected* saving because it was based on historical usage data as well as the (conservative) assumption of no behavioral response to changes in price. This information was provided to all the customers in this experiment.

[Figure 6 about here]

Figure 6 shows the distribution of customers in terms of their expected savings from dynamic pricing. The distribution implies that even with no behavioral response to price changes (zero price elasticity assumption), about half of customers can lower their payment simply by switching to the dynamic tariff. This segment of customers are called structural winners (Borenstein, 2013). Our description in Section 2.2 implies that theoretically, customers in this segment may be more likely to adopt dynamic pricing through *selection on level*. Customers at the other end of the scale are called structural losers. These customers use more electricity in peak hours than in off-peak hours, meaning that they are likely to lose money from dynamic pricing unless they are price-elastic and can re-optimize their usage in response to changes in price.

4.3   Timeline of the Experiment

The experiment consisted of two stages. The first stage was each customer's decision to adopt dynamic pricing. In June 2014, customers were notified that they were eligible to switch from default flat pricing to dynamic pricing. Customers received a graphical explanation of the two electricity tariffs, as presented in Figure A.1. In addition, we explicitly presented each customer's annual expected saving from dynamic pricing based on his/her historical hourly usage.

The second stage of the experiment was the implementation of dynamic pricing for customers who decided to switch. Customers who switched and those who stayed with the default plan received the same type of daily information, including the information about their usage and price, via the in-home display and text messages.

5   Empirical Analysis

In this section, we use the data from our RCT to estimate the selection equation, MTE, and MTRs, which are key parameters for the welfare analysis in Section 6.

5.1   Descriptive Evidence on Selection

We first present descriptive evidence on selection. Figure 7 shows that 31% of the the baseline group and 48% of the incentive group, which received a $60 take-up incentive, adopted dynamic pricing. A key question is whether consumers self-selected based on their expected savings from dynamic pricing, that is, whether there is evidence of selection on the *level* (Einav et al., 2013).

[Figure 7 about here]

We test this prediction in Figure 8. It shows the histogram of the expected savings from dynamic pricing and frequency of take-up. The baseline group's figure shows evidence of selection on the level. Consumers with positive expected savings were more likely to adopt. Moreover, the take-up incentive group's figure indicates that the $60 take-up incentive had different impacts on take-up between structural winners (i.e., customers with positive expected savings) and structural losers (i.e., customers with negative expected savings). This result is consistent with the fact that the $60 take-up incentive is likely to be pivotal for structural losers, whereas it may not be pivotal for structural winners, who would adopt dynamic pricing without the take-up incentive.

[Figure 8 about here]

Figure 9 summarizes this different impact of the take-up incentive by presenting the take-up rate of each group over the expected savings. The upward trend for the baseline group suggests evidence of selection on the level. For the take-up incentive group, we observe a horizontal shift to the left compared with the baseline group. The shift is approximately $60, which is consistent with the fact that they received a $60 take-up incentive.

[Figure 9 about here]

The figure also shows that the effect of the take-up incentive ($Z = \$60$) on take-up (i.e., the vertical distance between the two lines) is substantially different across customers with different

expected savings. This variation is important for our estimation of the MTE and MTRs in the next section. In econometrics terminology, this figure implies that although instrument Z is discrete, the effect of Z on the take-up of treatment D is substantially different across different values of observables X, which in this case are the expected savings from dynamic pricing. We apply this variation to the estimation method developed by Brinch, Mogstad and Wiswall (2017) to estimate the flexible functional forms of the MTE and MTRs in Section 5.3.

With these pieces of descriptive evidence on selection in mind, we provide a formal statistical analysis of the selection equation in the next subsection.

5.2 Selection Equation

Recall that in equation (2), we model the selection equation by $D = 1$ if $S_1 - S_0 = \nu(X, Z) - V > 0$, where $S_1 - S_0$ is the net surplus of adopting dynamic pricing, $\nu(X, Z)$ is the observable part of the selection equation with a flexible function $\nu(.)$, observables X, and instrument Z. V is the unobserved disutility for adopting dynamic pricing with distribution $F_V(.)$.

A common way to estimate the selection equation is to assume a certain parametric distribution for $F_V(.)$, such as the logistic and normal distributions. A disadvantage of such methods is that it is not guaranteed that such a distributional assumption is valid for unobservables V. For this reason, we adopt the semi-nonparametric approach developed by Gallant and Nychka (1987). This approach approximates the unknown density (f_V) using a Hermite polynomial expansion of the form $f_V(v) = \frac{1}{\psi} g(v)^2 \phi(v)$, where $\phi(\cdot)$ is the standardized Gaussian density, $g(v) = \sum_{k=0}^{K} \gamma_k v^k$ is a polynomial in v of order of K, and $\psi = \int_{-\infty}^{\infty} g(v)^2 \phi(v) dv$ is a normalization factor that ensures f is a probability distribution function.

This semi-nonparametric approach is in a sense still parametric because it fits the data to a flexible parametric distribution $\frac{1}{\psi} g(v)^2 \phi(v)$. As shown by Gallant and Nychka (1987), this distribution is sufficiently flexible for approximating various continuous distributions. Although the semi-nonparametric approach is less flexible than fully non-parametric approaches, it facilitates our counterfactual analysis because we can recover the distribution of $f_V(.)$.

Table 3 shows the estimation results for the selection equation. We include instrument Z—the randomly assigned dummy variable for eligibility for the take-up incentive of \$60—and observables X in the estimation. We have several demographic characteristics from our pre-experimental survey.

In addition, we include the two risk preference variables, risk aversion and certainty premium, which were obtained from the risk solicitation method in Section 4.2.

[Table 3 about here]

Table 3 suggests that the take-up incentive increased the take-up rate. In addition, the risk aversion, certainty premium, and employment status are negatively associated with selection. Years of schooling and expected savings from dynamic pricing are positively associated with selection. In the last two columns, we include dummy variables for the deciles of expected savings instead of estimating the linear relationship. These results are consistent with those in Figure 9.

The table also shows the estimated coefficients for γ_k in the polynomial $g(v) = \sum_{k=0}^{K} \gamma_k v^k$ in density $f_V(v) = \frac{1}{\psi} g(v)^2 \phi(v)$. Through cross-validation, we find that the optimal order of the polynomial is $K = 3$. The chi statistic for the likelihood ratio test of the standard normal distribution against the estimated $f_V(v)$ is 54.74, with a p-value of 1.37e-13. Thus, we reject the hypothesis that the distribution is normal.

Advantage of having the take-up incentive $Z = \$60$ is that it allows us to calculate the monetary values of the estimated coefficients. We denote the estimated coefficient for Z by $\hat{\alpha}$. Then, we can estimate the monetized value of any other coefficient, say $\hat{\beta}$, by $\hat{\beta} \times (60/\hat{\alpha})$. In addition, the the monetized value of the net surplus can be obtained by $(\hat{\nu}(X, Z) - V) \times (60/\hat{\alpha})$.

5.3 Marginal Treatment Effects and Marginal Treatment Responses

In addition to the selection equation, the other key inputs for the framework in Section 3 are the MTE and MTRs. We use the method developed by Brinch, Mogstad and Wiswall (2017) to estimate the MTE and MTRs with a discrete instrument (i.e., take-up incentive Z), which was randomly assigned in our field experiment.

We model potential outcomes for electricity usage at time t for tariff choice $j = \{0, 1\}$ by $Y_{t,j} = \mu_{t,j}(X) + U_{t,j}$. $\mu_{t,j}(X)$ are any flexible functions of observables X, and $U_{t,j}$ are the unobserved part of the potential outcomes. Using this equation, the MTRs can be written as $m_{t,j}(x, u) = E[Y_{t,j}|X = x, U = u] = \mu_{t,j}(x) + E[U_{t,j}|X = x, U = u]$. Recall that $U_{t,j}$ are the unobserved part of the potential outcomes, whereas U is the unobserved part of the selection equation in Section 3.

Using a discrete instrument, we can estimate the MTE and MTRs by assuming that they are linear in U. However, the relationship between these parameters and U can be nonlinear, which would provide important implications for the link between the selection and potential outcomes. To address this point, Brinch, Mogstad and Wiswall (2017) suggest that having the following separability assumption allows the MTE and MTRs to be nonlinear in U:

$$m_{t,j}(x, u) = E[Y_{t,j}|X = x, U = u] = \mu_{t,j}(x) + E[U_{t,j}|U = u], \tag{12}$$

for $j \in \{0, 1\}$. This implies that the MTE can be written by

$$Y_t^{MTE}(x, u) = m_{t,1}(x, u) - m_{t,0}(x, u) = \mu_{t,1}(x) - \mu_{t,0}(x) + E[U_{t,1} - U_{t,0}|U = u]$$

$$= \mu_t(x) + E[U_{t,1} - U_{t,0}|U = u], \tag{13}$$

where $\mu_t(x) \equiv \mu_{t,1}(x) - \mu_{t,0}(x)$. It is useful to summarize the advantages and limitations of the separability assumption. First, as mentioned above, the separability assumption allows us to estimate the MTE and MTRs as nonlinear functions of U. This is important because it allows the selection effects to enter the potential outcomes and treatment effects in a flexible way. Second, even with the separability assumption, the equations above are flexible functions of both the observables ad unobservables. Third, the separability assumption imposes $E[U_{t,j}|X = x, U = u] = E[U_{t,j}|U = u]$. That is, the conditional expectation of the unobserved parts of the potential outcomes, conditional on the unobserved part of the selection equation, does not depend on observables. For example, consider two consumers who have the same value of unobservables in the selection equation but have different values of observables. The separability assumption imposes that the expected value of $U_{t,j}$ is the same for these example customers.

The separability assumption allows us to estimate the flexible functions of $E[U_{t,1} - U_{t,0}|U = u]$ for the following reason. Suppose that the impacts of the discrete instrument (Z) on treatment take-up (D) are heterogeneous among individuals with different values of X. Figure 9 presents an empirical example of this variation. Our take-up incentive (Z) had different impacts on D for structural winners (i.e., those who had higher values of X) and structural losers (i.e., those who had lower values of X). With the separability assumption, we can thus exploit this variation to

estimate the flexible functions of the MTE and MTRs.

Our estimation procedure follows the two-step approach of Brinch, Mogstad and Wiswall (2017). First, we obtain estimates for propensity score $\hat{P}(X, Z)$ from the selection equation in Table 3. Second, we estimate $\mu_{t,0}(X)$, $\mu_{t,1}(X)$, $E[U_{t,0}|U = u]$, and $E[U_{t,1}|U = u]$ using the double residual regression method of Robinson (1988), as modified by Heckman, Ichimura and Todd (1997). This method first estimates the relationship between potential outcomes and covariates (i.e.,$\mu_{t,0}(X)$ and $\mu_{t,1}(X)$) and then estimates $E[U_{t,0}|U = u]$ and $E[U_{t,1}|U = u]$ semi-parametrically by running local quadratic regression of $Y_{t,1} - \hat{\mu}_{t,1}(X)$ and $Y_{t,0} - \hat{\mu}_{t,0}(X)$ on $\hat{P}(X, Z)$. This method allows us to estimate $Y_t^{MTE}(x, u) = \mu_t(x) + E[U_{t,1} - U_{t,0}|U = u]$. As noted by Brinch, Mogstad and Wiswall (2017), this method allows semi-parametric estimation in the range of U that has sufficient overwrap of the propensity scores from the treated and untreated customers. In our case, this range is $U \in [0.08, 0.91]$.

Before we investigate unobserved heterogeneity in the MTE, we examine observed heterogeneity described by the estimates of $\mu_t(x)$. Table 4 shows the estimates of $\mu_t(x)$ for peak hours, where treated customers faced an increase in their electricity price. The negative treatment effect thus implies a reduction in usage in response to this higher price. The positive coefficient on the expected savings implies that consumers with larger expected savings (i.e., structural winners) had *lower* reductions in usage in response to a price increase. Recall that the selection equation in Table 3 and Figure 8 show evidence of selection on the *level*; customers with larger expected savings were more likely to adopt dynamic pricing. However, the result in Table 4 implies that structural winners are likely to reduce electricity usage to a lower extent, suggesting that the welfare gains from structural winners are smaller than those from structural losers. This finding implies that encouraging structural losers to adopt dynamic pricing (e.g., by providing relatively larger take-up incentives) is likely to be welfare-improving.

[Table 4 about here]

In addition to expected savings, similar phenomena can be observed for risk aversion and certainty premiums. Table 3 suggests that customers with higher risk aversion and certainty premiums are less likely to adopt dynamic pricing. However, Table 4 indicates that these customers produce a larger reduction in peak-hour usage, leading to larger welfare gains. Therefore, customers with

higher risk aversion and certainty premiums are observable types that may benefit from relatively large take-up incentives when the take-up incentive can be differentiated by the observables.

On the contrary, selection on years of schooling and employment status is likely to be in line with the social planner's objective. Table 3 indicates that customers with more years of schooling and those unemployed are more likely to adopt dynamic pricing, and Table 4 suggests that these customers would reduce peak-hour usage to a larger degree. This finding suggests positive selection on the *slope* based on these observables. Customers with these observable types such as highly educated consumers are more price-elastic and also more likely to self-select into dynamic pricing.[7]

The question is whether there is selection on the slope through *unobserved* heterogeneity, conditional on selection through the observables. The advantage of the MTE in our context is that the relationship between the MTE and unobserved consumer type U, that is, $E[U_{t,1} - U_{t,0}|U = u]$ in the MTE function, provides a direct test for this question.

Figure 10 shows the estimated MTE function against consumer type U. Panels A and B show the results for peak-hour electricity usage in the summer and winter, respectively. In other words, the curvatures of the lines trace out the semi-parametric estimates of $E[U_{t,1} - U_{t,0}|U = u]$. These figures show evidence of selection on the slope through unobserved heterogeneity—consumers with lower U (i.e., those who are more likely to adopt dynamic pricing for unobserved reasons) have larger treatment effects (i.e., more reductions in electricity usage in response to the price increase). The treatment effect diminishes as we move to customers who have higher U, namely, those who are less likely to adopt dynamic pricing for unobserved reasons.

[Figure 10 about here]

We conduct the same analysis for off-peak hours and show the results in the appendix. In off-peak hours, those who selected into dynamic pricing had a small decrease in price (a change from 26 cents to 21 cents per kWh). In theory, they could thus increase their usage in off-peak hours, and this increase is welfare-improving as shown in Figure 1. However, we do not find strong evidence of such behavior in our experiment.[8] This finding is consistent with those of previous studies (Wolak,

---

[7]Table 4 also suggests that household income is not statistically significant conditional on the other observables included in the estimation.

[8]The MTE in the summer off-peak hours in the appendix shows suggestive evidence of selection on the slope in off-peak hours, but it is not statistically significant.

). A possible reason is that the change in off-peak price was small relative to the large change in peak-hour price, and thus, consumers may not have strongly reacted to such a small price change.

In sum, the selection on the slope found in this section provides two key implications for social welfare and take-up incentives. First, there is a chance that a take-up incentive can improve welfare if consumers' self-selection based on observables and unobservables results in take-up below the socially optimal level. Second, however, the marginal return from such an incentive is likely to diminish because consumers with higher U are likely to produce a smaller social welfare gain. In addition, consumers with higher U induce a higher utility loss when they are pushed into take-up because the definition of U implies that their unobserved disutility from adopting dynamic pricing is likely to be larger. This suggests that at some level of take-up incentive, the marginal net welfare gain can reach to zero. With this intuition, we investigate the welfare implications more formally in the next section by combining the empirical results in this section and framework in Section 3 .

6   Welfare Analysis

In this section, we examine the welfare implications of various policies based on the framework developed in Section 3 and parameter estimates from Section 5.

6.1   Welfare Gains with Respect to Take-Up Incentive (Z)

We first explore how a policy's social welfare gain per capita $W^{ITT}(x, z)$ changes with respect to take-up incentive Z. To perform this calculation, we plug the parameter estimates from our empirical analysis into equation (8). Our framework allows $W^{ITT}(x, z)$ to vary between individuals with different values of observables X. For illustration purpose, we use given values of X to show the following figures.

Figure 11 shows the welfare gain per capita with respect to the take-up incentives. The figure shows that providing a positive take-up incentive improves welfare compared with a policy with a zero take-up incentive. In addition, the welfare gain per capita is concave in the take-up incentives—it increases to a certain value of Z and then starts to decline.

This concavity occurs for two primary reasons. The first reason is the shape of the MTE estimated in the previous section. Higher Z induces more consumers to adopt dynamic pricing.

However, marginal consumers with higher Z do not provide as large treatment effects (i.e., changes in electricity usage in response to changes in price) as infra-marginal consumers who would adopt dynamic pricing with lower Z. The second reason is heterogeneity in the unobserved disutility for adopting dynamic pricing, which is term V in the selection equation. The selection equation implies that as we increase Z, consumers with higher V (i.e., those with higher disutility from adoption) would adopt dynamic pricing. This mechanism also contributes to the diminishing return of Z in the figure.

[Figure 11 about here]

Figure 12 also presents this point, showing the *marginal* welfare gain per capita with respect to the take-up incentive $\left(\frac{dW^{ITT}(x,z)}{dz} \right)$. The marginal welfare gain is high when Z is near zero but diminishes and eventually becomes negative as the benefit of marginal consumers becomes smaller than the cost.

[Figure 12 about here]

6.2 Welfare Comparison Between Counterfactual Policies

Table 5 compares the welfare implications of four policies—the two policies from the RCT ($Z = 0$ and $Z = 60$) and two counterfactual policies. The first counterfactual is a policy with an optimal uniform take-up incentive ($Z = z^*$) that maximizes the welfare gain per capita $W^{ITT}(x, z)$ under the constraint that a take-up incentive cannot be differentiated by observables X. The second counterfactual is a policy with optimal differentiated take-up incentives ($Z = z^*(x)$) that maximize $W^{ITT}(x, z)$ with take-up incentives that can vary by observables X.

The outcome variables of interest are 1) the take-up rate, 2) the ATE on the treated on the welfare gain (W^{ATET}), and 3) the intention-to-treat on the welfare gain (W^{ITT}). In general, the ultimate policy goal is to maximize W^{ITT} because it measures the policy's overall welfare gain per capita. W^{ATET} is also an informative outcome variable for understanding the welfare gain per capita among consumers who adopt dynamic pricing in different policy scenarios.

[Table 5 about here]

Rows 1 and 2 in Table 5 are obtained from the experimental variation in our RCT. Providing a take-up incentive ($Z = 60$) increases the take-up rate from 31% to 48%. At the same time, providing

such an incentive reduces W^{ATET} because the marginal welfare gain decreases in Z, as described in Section 6.1. Overall, providing $Z = 60$ increases W^{ITT} from \$18.1 to \$23.7 per consumer per year.

The results in rows 3 and 4 are obtained by counterfactual policy simulations based on the model described in Section 3 and parameter estimates from Section 5. Row 3 investigates the case with the optimal uniform take-up incentive ($Z = z^*$). The optimal uniform incentive is \$98.1 and increases the take-up rate to 65%. Compared with the case with $Z = 60$, $Z = z^*$ improves W^{ITT}, but the additional welfare gain is not large. This is because W^{ATET} diminishes with Z, as shown in column 4.

Row 4 shows the policy outcomes with the optimal differentiated take-up incentive ($Z = z^*(x)$). This policy increases the take-up rate to 66%. Compared with the case with $Z = z^*$, $Z = z^*(x)$ further improves W^{ITT} to \$31.5 because the differentiated take-up incentives can target consumer types that generate high W^{ATET} based on observables X.[9] For example, suppose that consumers with certain values of observables have high W^{MTE} for a wider range of U than other consumers. Then, an optimal differentiated incentive can induce these consumers with a higher subsidy specific to this observable type. Another example is consumers with certain values of the observables who have high W^{MTE} but a low propensity score. If incentives can be differentiated by the observables, policymakers can capture these customers by increasing the incentive for this observable type.

This point is mathematically seen in equation (10) in Section 3. The optimal incentive formula exploits both the propensity score from the selection equation ($P(x, z)$) and the MTE of the welfare gain ($W^{MTE}(x, z, u)$). The results in this table confirm that the optimal subsidies indeed improve welfare empirically with our data.

6.3 Welfare Comparison with a Mandate

As discussed in the Introduction, the vast majority of countries rely on voluntary take-up policies for the adoption of dynamic electricity pricing because a mandate is politically infeasible. However, it is still useful to discuss what conditions could make a mandate more welfare-enhancing than the counterfactual policies discussed above.

While our study cannot comprehensively analyze a mandate, we can use our model to provide a discussion below, which would be useful for comparing a mandate and counterfactual policies

---

[9]The 25th, 50th, and 75th percentiles of the differentiated take-up incentives are \$50.2, \$86.6, and \$123.1.

considered in the previous section.

To analyze a mandate, a key question is how it would change consumers' utility, particularly their disutility for adopting dynamic pricing. For example, consumers may dislike adopting dynamic pricing because they have an unobserved preference for non-dynamic pricing relative to dynamic pricing. Another example is that they may have inertia or switching costs for moving from default pricing to dynamic pricing. An important question is how a mandate affects these factors.

We can consider two possibilities. The first is that a mandate does not change or may increase the costs of these factors. In this case, consumers have the same or higher disutility from these factors as they have under voluntary take-up policies. In this case, our results suggest that a mandate is likely to be inferior to optimal take-up incentive policies because additional take-up beyond the optimal take-up in Table 5 lowers the welfare gain (W^{ITT}). This is because the social welfare gain from the marginal consumer is lower than the welfare loss from the consumer's disutility for adopting dynamic pricing.

The second possibility is that a mandate may lower consumers' disutility from adopting dynamic pricing. For example, a mandate may be able to lower switching costs. For this possibility, we can use our counterfactual simulation to examine how much such a benefit has to be to make a mandate policy more welfare-enhancing than other policies. In our counterfactual simulation, if we increase the take-up rate to 100%, the welfare gain per capita (W^{ITT})—note that this calculation still includes consumers' disutility from adoption—becomes $5.5 per year per consumer. This implies that a mandate can be more welfare-improving than the take-up incentive policy if it can reduce consumers' disutility by more than $26 ($= 31.5 - 5.5$) per customer per year. This result implies that if the net benefit of a mandate (the economic benefit minus the political cost) exceeds this number, a mandate can be a preferred policy than voluntary take-up policies.

7 Conclusion

In this study, we investigate a problem in which policymakers need to screen self-selected individuals by unobserved heterogeneity in social welfare gains from a policy intervention. In our framework, the marginal treatment effects and marginal treatment responses arise as key statistics to characterize social welfare. We apply this framework to a randomized field experiment on electricity plan choice.

27

In the experiment, consumers were offered socially efficient dynamic pricing with randomly assigned financial take-up incentives. We find that price-elastic consumers, who generate larger social welfare gains, are more likely to self-select. Finally, we use counterfactual simulations to quantify the optimal take-up incentives by exploiting observed and unobserved heterogeneity in selection and welfare gains.
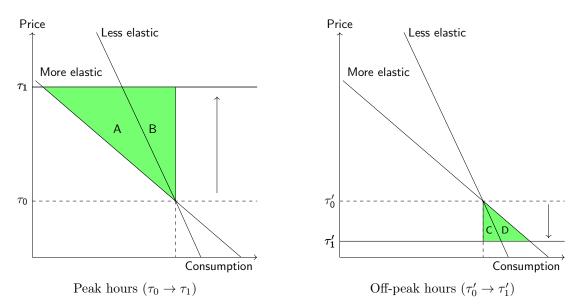
References

Allcott, Hunt, and Michael Greenstone. 2017. "Measuring the Welfare Effects of Residential Energy Efficiency Programs." NBER Working Paper 23386.

Borenstein, Severin. 2002. "The Trouble With Electricity Markets: Understanding California's Restructuring Disaster." *Journal of Economic Perspectives*, 16(1): 191–211.

Borenstein, Severin. 2013. "Effective and Equitable Adoption of Opt-In Residential Dynamic Electricity Pricing." *Review of Industrial Organization*, 42(2): 127–160.

Borenstein, Severin, and Lucas W. Davis. 2016. "The Distributional Effects of US Clean Energy Tax Credits." *Tax Policy and the Economy*, 30(1): 191–234.

Brinch, Christian N., Magne Mogstad, and Matthew Wiswall. 2017. "Beyond LATE with a Discrete Instrument." *Journal of Political Economy*, 125(4): 985–1039.

Callen, Michael, Mohammad Isaqzadeh, James D. Long, and Charles Sprenger. 2014. "Violence and Risk Preference: Experimental Evidence from Afghanistan." *American Economic Review*, 104(1): 123–48.

Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics*, 1(1): 451–488.

Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg. 2018. "Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance." *Journal of Political Economy*, 126(6): 2356–2409.

Deshpande, Manasi, and Yue Li. 2019. "Who Is Screened Out? Application Costs and the Targeting of Disability Programs." *American Economic Journal: Economic Policy*, 11(4): 213–248.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf. 2010. "Optimal Mandates and the Welfare Cost of Asymmetric Information: Evidence From the U.K. Annuity Market." *Econometrica*, 78(3): 1031–1092.

Einav, Liran, Amy Finkelstein, Stephen Ryan, Paul Schrimpf, and Mark R. Cullen. 2013. "Selection on Moral Hazard in Health Insurance." *American Economic Review*, 103(1): 178–219.

Einav, Liran, Amy Finkelstein, Yunan Ji, and Neale Mahoney. 2020. "Voluntary Regulation: Evidence from Medicare Payment Reform." NBER Working Paper 27223.

Eisenhauer, Philipp, James J. Heckman, and Edward Vytlacil. 2015. "The Generalized Roy Model and the Cost-Benefit Analysis of Social Programs." *Journal of Political Economy*, 123(2): 413–443.

Faruqui, Ahmad, and Sanem Sergici. 2011. "Dynamic Pricing of Electricity in the Mid-Atlantic Region: Econometric Results from the Baltimore Gas and Electric Company Experiment." *Journal of Regulatory Economics*, 40(1): 82–109.

FERC. 2011. "2010 Assessment of Demand Response and Advanced Metering." *Staff Report.*

Finkelstein, Amy, and Matthew J Notowidigdo. 2019. "Take-up and Targeting: Experimental Evidence from SNAP." *Quarterly Journal of Economics*, 134(3): 1505–1556.

Fowlie, Meredith, Catherine Wolfram, Anna Spurlock, Annika Todd, Patrick Baylis, and Peter Cappers. 2017. "Default Bias and Follow-On Behavior: Evidence from an Electricity Pricing Program." NBER Working Paper 23553.

Gallant, A., and Douglas W. Nychka. 1987. "Semi-nonparametric Maximum Likelihood Estimation." *Econometrica*, 55(2): 363–90.

Heckman, James J. 2010. "Building Bridges between Structural and Program Evaluation Approaches to Evaluating Policy." *Journal of Economic Literature*, 48(2): 356–398.

Heckman, James J., and Edward J. Vytlacil. 1999. "Local Instrumental Variables and latent Variable Models for Identifying and Bounding Treatment Effects." *Proceedings of the National Academy of Sciences of the United States of America*, 96(8): 4730–4734.

Heckman, James J., and Edward Vytlacil. 2001. "Policy-Relevant Treatment Effects." *The American Economic Review*, 91(2): 107–111.

Heckman, James J., and Edward Vytlacil. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica*, 73(3): 669–738.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. "How Details Make a Difference: Semiparametric Estimation of the Partially Linear Regres- sion Model." *Manuscript.*

Hortaçsu, Ali, Seyed Ali Madanizadeh, and Steven L. Puller. 2017. "Power to Choose? An Analysis of Consumer Inertia in the Residential Electricity Market." *American Economic Journal: Economic Policy*, 9(4): 192–226.

Ito, Koichiro, Takanori Ida, and Makoto Tanaka. 2017. "Information Frictions, Inertia, and Selection on Elasticity: A Field Experiment on Electricity Tariff Choice." *Working Paper.*

Ito, Koichiro, Takanori Ida, and Makoto Tanaka. 2018. "Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand." *American Economic Journal: Economic Policy*, 10(1): 240–67.

Joskow, Paul L. 2012. "Creating a Smarter U.S. Electricity Grid." *Journal of Economic Perspectives*, 26(1): 29–48.

Joskow, Paul L., and Catherine D. Wolfram. 2012. "Dynamic Pricing of Electricity." *American Economic Review*, 102(3): 381–85.
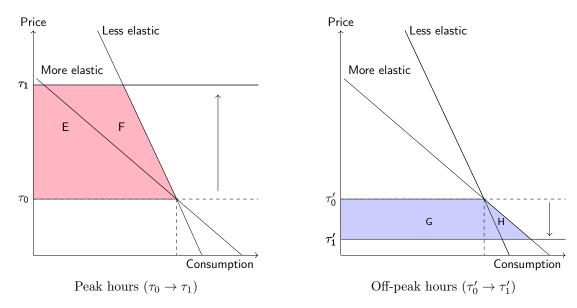
Kitagawa, Toru, and Aleksey Tetenov. 2018. "Who Should Be Treated? Empirical Welfare Maximization Methods for Treatment Choice." *Econometrica*, 86(2): 591–616.

Kleven, Henrik J. 2020. "Sufficient Statistics Revisited." NBER Working Paper 27242.

Kline, Patrick, and Christopher R. Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start." *The Quarterly Journal of Economics*, 131(4): 1795–1848.

LaLonde, Robert J. 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *The American Economic Review*, 76(4): 604–620.

Marschak, Jacob. 1953. *Economic Measurements for Policy and Prediction.* Cowles Comm. Res. Econ., Monogr. No. 14. New York: Wiley.

Mogstad, Magne, Andres Santos, and Alexander Torgovitsky. 2018. "Using Instrumental Variables for Inference About Policy Relevant Treatment Parameters." *Econometrica*, 86(5): 1589–1619.

Robinson, P. M. 1988. "Root-N-Consistent Semiparametric Regression." *Econometrica*, 56(4): 931–954.

Wolak, Frank A. 2006. "Residential customer response to real-time pricing: the Anaheim critical-peak pricing experiment." *Working Paper.*

Wolak, Frank A. 2011. "Do Residential Customers Respond to Hourly Prices? Evidence from a Dynamic Pricing Experiment." *The American Economic Review*, 101(3): 83–87.

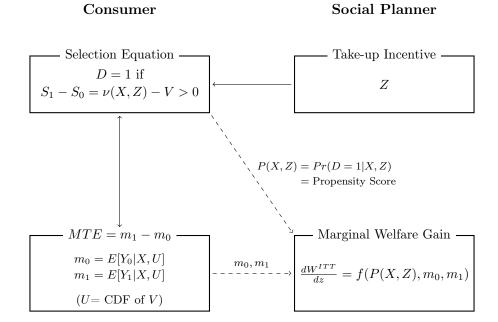Figure 1: Social Welfare Gains from Dynamic Pricing

Peak hours ($\tau_0 \to \tau_1$) Off-peak hours ($\tau_0' \to \tau_1'$)

Note: These figures show welfare gains from dynamic pricing for two hypothetical demand curves. Consider a change from flat pricing (τ_0, τ_0') to dynamic pricing (τ_1, τ_1'), which reflects the time-varying marginal cost of electricity. When the more-elastic consumer adopts dynamic pricing, the welfare gain (i.e., the reduction in deadweight loss) is $A + B$ in peak hours and $C + D$ in off-peak hours. The welfare gain from the less-elastic consumer is A in peak hours and C in off-peak hours. Thus, the more-elastic consumer produces a larger welfare gain upon take-up. These figures do not include the inertia/switching costs from adopting dynamic pricing, which we incorporate in our model and estimation.

Figure 2: Consumers' Private Gains from Dynamic Pricing

Peak hours $(\tau_0 \rightarrow \tau_1)$ Off-peak hours $(\tau'_0 \rightarrow \tau'_1)$

Note: These figures show consumers' private gains from dynamic pricing for two hypothetical demand curves. When the more-elastic consumer adopts dynamic pricing, the private gain (i.e., the change in the consumer surplus) is $-E$ in peak hours and $G + H$ in off-peak hours. The private gain for the less-elastic consumer is $-(E + F)$ in peak hours and G in off-peak hours. Thus, the more-elastic consumer has a larger private gain upon take-up. These figures do not include inertia/switching cost from adopting dynamic pricing, which we incorporate in our model and estimation.
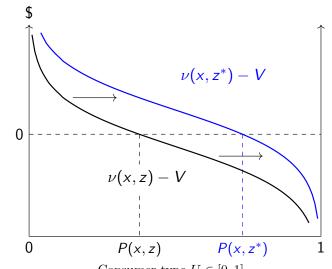
Figure 3: Conceptual Framework: Overview

Consumer   **Social Planner**

Selection Equation
$$D = 1 \text{ if}$$
$$S_1 - S_0 = \nu(X, Z) - V > 0$$

Take-up Incentive
$$Z$$

$$P(X, Z) = Pr(D = 1 | X, Z)$$
$$= \text{Propensity Score}$$

$MTE = m_1 - m_0$
$$m_0 = E[Y_0 | X, U]$$
$$m_1 = E[Y_1 | X, U]$$
$$(U = \text{CDF of } V)$$

m_0, m_1

Marginal Welfare Gain
$$\frac{dW^{ITT}}{dz} = f(P(X, Z), m_0, m_1)$$

Note: This figure visualizes the conceptual framework in Section 3. The social planner provides financial take-up incentive Z for consumers to switch from inefficient time-invariant pricing to efficient dynamic pricing. A consumer's switching decision is described by the Roy model with a selection equation and potential outcomes. In Section 3.1, we use this framework to derive the MTE and MTRs (m_0 and m_1 in the diagram). We then show the connection between the consumer's problem and social planner's problem in Section 3.2. A key insight is that the social planner's objective function (i.e., the welfare gain from a policy) can be written as a function of the MTE, MTRs, and propensity score from the selection equation.
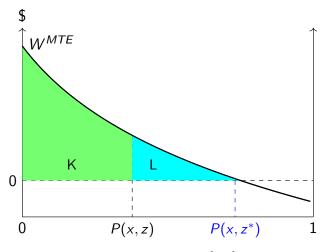
Figure 4: Conceptual Framework: Consumer's Selection and Social Planner's Welfare Gains

A) Consumer: Selection Equation: $D = 1\{\nu(x, z) - V > 0\}$

Consumer type $U \in [0, 1]$
(unobserved disutility for adopting dynamic pricing)
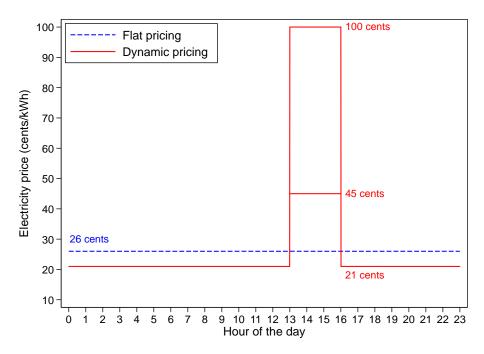
B) Social Planner: Marginal Welfare Gains (W^{MTE})

Consumer type $U \in [0, 1]$
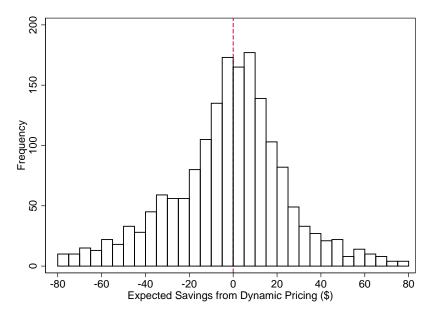(unobserved disutility for adopting dynamic pricing)

Note: These figures visualize the link between a consumer's selection and the social planner's welfare gains from the consumer. See the text in Section 3.3. Panel A illustrates a consumer's hypothetical selection equation $S_1 - S_0 = \nu(x, z) - V$ over unobserved consumer type U. It is downward-sloping because U is defined as the disutility for adopting dynamic pricing. With the initial level of $Z = z$, consumers with $U \leq P(x, z)$ self-select, where $P(x, z)$ is the propensity score with observables $X = x$ and the take-up incentive $Z = z$. Take-up increases to $P(x, z^*)$ if the take-up incentive is increased to $Z = z^*$. Panel B illustrates a hypothetical function of $W^{MTE}(x, z, u)$. In this example, the welfare gain per capita, $W^{ITT}(x, z)$, equals to K. This can be further increased to $K + L$ by increasing the take-up incentive to $Z = z^*$. The actual slope of W^{MTE} can be upward, downward, or flat, and we empirically estimate it in Section 5.

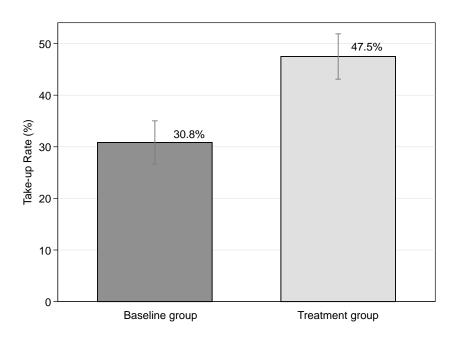Figure 5: Electricity Price Plans

Note: This figure shows the electricity tariffs for the default flat price plan and new dynamic pricing. The dynamic pricing plan charges 21 cents/kWh in off-peak hours and either 45 or 100 cents/kWh in peak hours depending on the day. Peak hours are between 1 pm and 4 pm in summer and between 5 pm and 8 pm in winter.

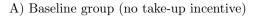Figure 6: Expected Savings from Dynamic Pricing

Note: This figure shows the histogram of the expected savings from dynamic pricing. The expected savings are calculated based on each customer's past usage, assuming no behavioral response to changes in price.
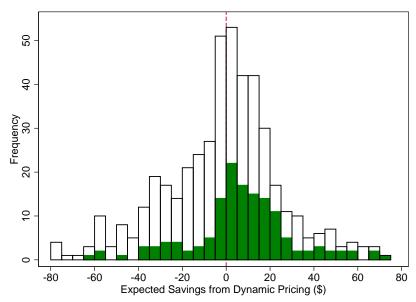
Figure 7: Take-Up Rates of Dynamic Pricing
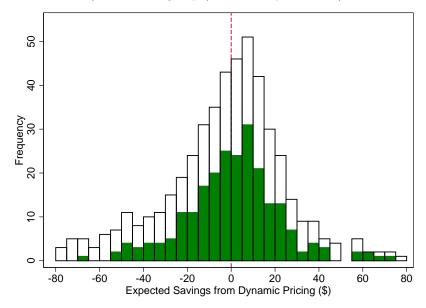
Note: The figure shows the take-up rates for the baseline group and treatment group that received the take-up incentive. The bars are 95% confidence intervals.

Figure 8: Selection on the *Level*: Take-Up Rates and Expected Savings from Dynamic Pricing

A) Baseline group (no take-up incentive)
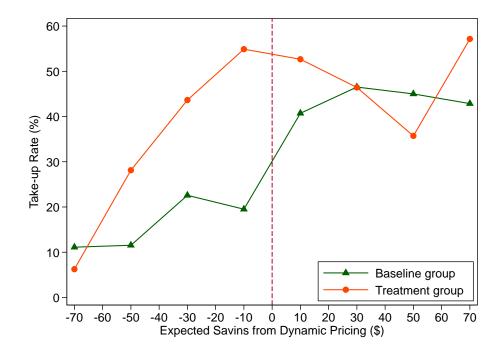
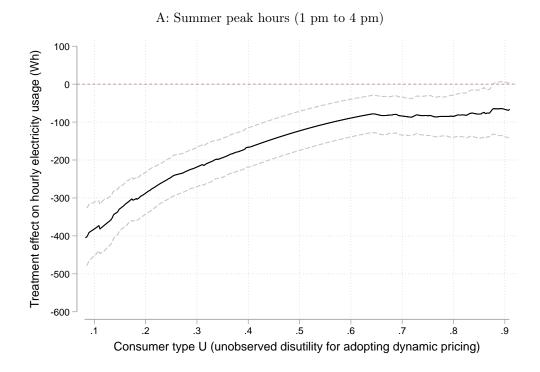B) Treatment group (with take-up incentive)

Note: These histograms show the distributions of customers in terms of their expected savings from dynamic pricing. The dark color indicates those who self-selected into dynamic pricing. The expected savings were calculated based on each customer's past usage, assuming no behavioral response to changes in price. Panel A indicates that structural winners, whose expected savings were larger, are more likely to self-select. Panel B indicates that the take-up incentive increased take-ups of structural losers, whose expected savings are small or negative.
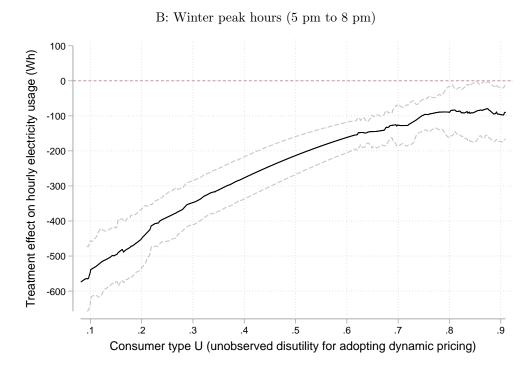
Figure 9: Heterogeneous Effects of the Take-Up Incentive on Take-up Rates

Note: This figure shows the take-up rate for each group by the expected savings from dynamic pricing. The take-up incentive caused different impacts on the take-up rates of structural winners and losers.

Figure 10: Selection on the *Slope*: Marginal Treatment Effects on Electricity Usage (Y^{MTE})

A: Summer peak hours (1 pm to 4 pm)

B: Winter peak hours (5 pm to 8 pm)

Note: This figure shows $Y_t^{MTE}(x, u)$ for peak hours in the summer and winter. We estimate equation (13) using hourly electricity usage data at the customer level and the estimation method developed by Brinch, Mogstad and Wiswall (2017). We allow the MTE to vary by X, and in these figures, we show the result for a given value of X. Consumers whose U is lower (i.e., those more likely to adopt dynamic pricing for unobserved reasons) have larger reductions in electricity usage in response to the increase in peak-hour price. The sample size of hourly electricity data in peak hours is 247,350 in the summer and 180,420 in the winter. The dashed lines show the 95 percent confidence intervals based on bootstrapped standard errors clustered at the customer level.

Figure 11: Welfare Gain per Capita: $W^{ITT}(x, z)$

Note: This figure shows the welfare gain per capita $W^{ITT}(x, z)$ with respect to different values of take-up incentive $Z = z$. We allow the $W^{ITT}(x, z)$ to vary by X, and in this figure, we show the result for a given value of X. For this observable type X, the estimated optimal take-up incentive is $85.
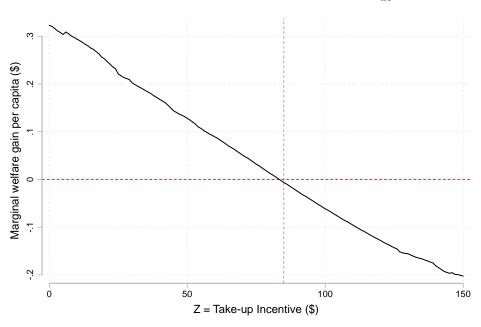
Figure 12: Marginal Welfare Gain per Capita: $\frac{dW^{ITT}}{dz}$

Note: This figure shows the marginal welfare gain per capita with respect to take-up incentive Z: $dW^{ITT}(x, z)/dz$. We allow $dW^{ITT}(x, z)/dz$ to vary by X, and in this figure, we show the result for a given value of X.

Table 1: Experimental Design

Group	Eligible to adopt dynamic pricing	Information provision	Take-up incentive
Baseline group	✓	✓	
Treatment group	✓	✓	✓

Table 2: Summary Statistics

	Baseline group	Treatment group
Household income (JPY10,000)	768.18 (343.46)	755.97 (334.23)
Square Meters	111.44 (31.09)	107.51 (30.64)
Number of Bedrooms	3.84 (1.15)	3.76 (1.09)
Age of Building	15.59 (11.68)	15.14 (11.39)
Number of room AC	3.17 (1.33)	3.10 (1.34)
Number of TV	2.12 (1.05)	2.07 (1.10)
Number of refrigerator	1.13 (0.38)	1.12 (0.37)
Electricity usage (kWh/day) in the pre-experimental period	13.09 (5.58)	13.02 (5.71)

Notes: This table shows the summary statistics of the two groups. Standard deviations are in parentheses. The sample size of customers is 468 for the baseline group and 502 for the treatment group. The sample size of hourly electricity usage data during the intervention period (147 days) is 1,651,104 for the baseline group and 1,771,056 for the treatment group.

Table 3: Selection Equation

Dependent Variable: $D_i = 1$(household i selected into dynamic pricing)

	(1)	(2)	(3)	(4)	(5)
Take-up incentive (60 USD)	0.704	0.636	0.640	1.248	1.013
	(0.168)	(0.155)	(0.160)	(0.234)	(0.236)
Risk aversion	-0.997	-0.967	-0.962	-1.472	-1.311
	(0.212)	(0.272)	(0.285)	(0.378)	(0.335)
Certainty premium	-0.986	-1.047	-1.065	-1.835	-1.248
	(0.326)	(0.334)	(0.356)	(0.538)	(0.455)
Expected saving (USD)	0.016	0.015	0.015		
	(0.003)	(0.004)	(0.004)		
Years of schooling		0.031	0.024		0.035
		(0.012)	(0.014)		(0.017)
Employed		-0.299	-0.423		-0.816
		(0.179)	(0.218)		(0.263)
Income (100,000 USD)			0.288		0.469
			(0.260)		(0.281)
Expected saving (decile 2)				0.886	0.366
				(0.299)	(0.298)
Expected saving (decile 3)				0.978	0.328
				(0.365)	(0.285)
Expected saving (decile 4)				0.976	0.431
				(0.338)	(0.263)
Expected saving (decile 5)				1.499	1.104
				(0.412)	(0.373)
Expected saving (decile 6)				2.714	3.211
				(0.497)	(0.391)
Expected saving (decile 7)				2.127	1.182
				(0.540)	(0.368)
Expected saving (decile 8)				1.976	1.377
				(0.419)	(0.355)
Expected saving (decile 9)				2.476	2.685
				(0.364)	(0.484)
Expected saving (decile 10)				1.846	1.296
				(0.392)	(0.372)
Parameters for density f_V					
γ_1	1.160	0.992	1.040	0.624	1.911
	(0.685)	(0.635)	(0.645)	(0.355)	(1.395)
γ_2	0.097	0.036	0.053	0.093	0.301
	(0.146)	(0.131)	(0.133)	(0.092)	(0.259)
γ_3	-0.451	-0.398	-0.408	-0.361	-0.669
	(0.194)	(0.177)	(0.183)	(0.103)	(0.380)
Log Likelihood	-563.65	-559.15	-558.48	-552.69	-545.72

Notes: This table shows the estimation results of the selection equation in equation (2). The sample size is 970. Standard errors are in parentheses.

Table 4: Observables in the Marginal Treatment Effect

Estimates of $\mu_t(x)$ in the MTE functions in equation (13)

	Summer		Winter	
	(1)	(2)	(3)	(4)
Risk aversion	-38.65	-39.95	-77.33	-77.06
	(24.88)	(24.88)	(29.51)	(29.54)
Certainty premium	-121.32	-122.57	-142.17	-141.91
	(30.52)	(30.50)	(36.19)	(36.21)
Expected saving (USD)	1.70	1.71	2.50	2.50
	(0.18)	(0.18)	(0.21)	(0.21)
Employed	52.05	63.08	54.93	52.62
	(13.86)	(15.26)	(16.43)	(18.12)
Years of schooling	-15.33	-14.22	-19.83	-20.07
	(2.65)	(2.73)	(3.15)	(3.24)
Income (100,000 USD)		-31.73		6.66
		(18.46)		(21.92)

Notes: This table shows the estimates of $\mu_t(x)$ in equation (13)—the observable part of the MTE function—for electricity usage in peak hours. The outcome variable is hourly electricity usage (in Wh) in peak hours. Treated customers had higher prices in peak hours and therefore the negative treatment effects imply reductions in electricity usage in response to increases in price. The sample size of hourly electricity data in peak hours is 247,350 in the summer and 180,420 in the winter. Bootstrapped standard errors clustered at the customer level are in parentheses.

Table 5: Welfare Comparison Between Counterfactual Policies

Policy	Evidence from	Take-up	Welfare gain: W^{ATET} ($/year/consumer)	Welfare gain: W^{ITT} ($/year/consumer)
$Z = 0$	RCT	31%	59.6	18.1
			(5.17)	(1.92)
$Z = 60$	RCT	48%	50.4	23.7
			(4.39)	(1.95)
$Z = z^*$	Counterfactual	65%	40.5	26.4
			(2.95)	(2.06)
$Z = z^*(x)$	Counterfactual	66%	47.7	31.5
			(1.86)	(1.82)

Notes: This table compares the welfare gains from four policies. The first two rows come from the results of our RCT. The last two rows are from counterfactual simulations based on the model presented in Section 3 and results in Section 5. $Z = z^*$ is the optimal uniform take-up incentive that does not depend on observables X, whereas $Z = z^*(x)$ is the optimal differentiated take-up incentive that is allowed to vary by observables X. Bootstrapped standard errors clustered at the customer level are in parentheses.

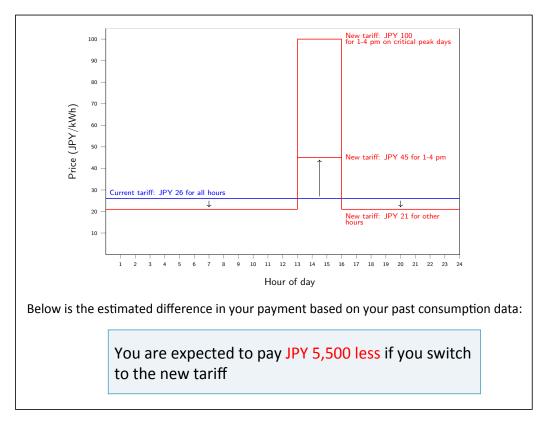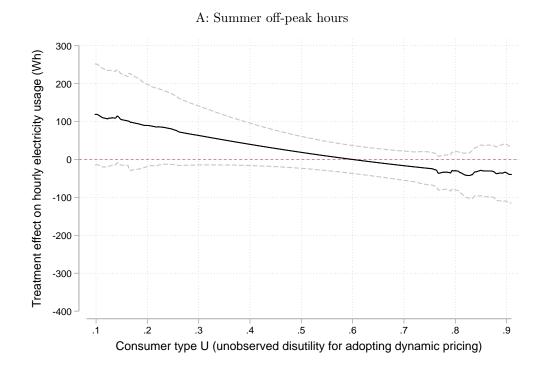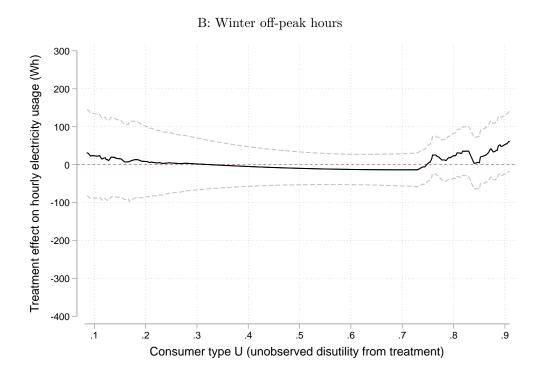# Online Appendix A: Additional Tables and Figures

Figure A.1: Information Provided to All Consumers

Note: This figure shows the information that provided to all consumers in our experiment. Customers were notified about the dynamic pricing structure and their expected savings.

Figure A.2: Off-peak hours: Marginal Treatment Effects on Electricity Usage (Y^{MTE})

A: Summer off-peak hours

B: Winter off-peak hours

Note: This figure shows $Y_t^{MTE}(x, u)$ for off-peak hours in the summer and winter. We estimate equation (13) using the estimation technique developed by Brinch, Mogstad and Wiswall (2017). We allow the MTE to vary by X, and in these figures, we show the result for a given value of X. The sample size of hourly electricity data in off-peak hours is 1,731,450 in the summer and 1,262,940 in the winter. The dashed lines show the 95 percent confidence intervals based on bootstrapped standard errors clustered at the customer level.

Table A.1: Elicitation of Risk Preference

Panel A: First set of questions to obtain q

q	Option A	Option B
0.1	$100	10% chance of $300, 90% chance of $0
0.2	$100	20% chance of $300, 80% chance of $0
0.3	$100	30% chance of $300, 70% chance of $0
0.4	$100	40% chance of $300, 60% chance of $0
0.5	$100	50% chance of $300, 50% chance of $0
0.6	$100	60% chance of $300, 40% chance of $0
0.7	$100	70% chance of $300, 30% chance of $0
0.8	$100	80% chance of $300, 20% chance of $0
0.9	$100	90% chance of $300, 10% chance of $0
1	$100	100% chance of $300, 0% chance of $0

Panel B: Second set of questions to obtain q'

q'	Option A	Option B
0.1	50% chance of $300, 50% chance of $0	10% chance of $300, 90% chance of $0
0.2	50% chance of $300, 50% chance of $0	20% chance of $300, 80% chance of $0
0.3	50% chance of $300, 50% chance of $0	30% chance of $300, 70% chance of $0
0.4	50% chance of $300, 50% chance of $0	40% chance of $300, 60% chance of $0
0.5	50% chance of $300, 50% chance of $0	50% chance of $300, 50% chance of $0
0.6	50% chance of $300, 50% chance of $0	60% chance of $300, 40% chance of $0
0.7	50% chance of $300, 50% chance of $0	70% chance of $300, 30% chance of $0
0.8	50% chance of $300, 50% chance of $0	80% chance of $300, 20% chance of $0
0.9	50% chance of $300, 50% chance of $0	90% chance of $300, 10% chance of $0
1	50% chance of $300, 50% chance of $0	100% chance of $300, 0% chance of $0

Notes: We asked customers to choose option A or B for each question. A customer's q and q' were obtained at which the choice between A and B was altered. Callen et al. (2014) show that q and q' can be used to calculate an individual's risk premium and certainty premium.

Table A.2: Experimental Sample and a Random Sample of Population in the Experimental Area

	Baseline group	Treatment group	Random sample of population
Household income (JPY10,000)	768.18 (343.46)	755.97 (334.23)	731.49 (435.46)
Square Meters	111.44 (31.09)	107.51 (30.64)	110.73 (45.95)
Age of Building	15.59 (11.68)	15.14 (11.39)	16.44 (9.08)
Number of room AC	3.17 (1.33)	3.10 (1.34)	3.48 (1.49)
Electricity usage (kWh/day)	13.09 (5.58)	13.02 (5.71)	12.28 (6.31)

Notes: This table shows summary statistics for the two groups in the experiment: the baseline group ($N = 468$) and the treatment group ($N = 502$) and a random sample of population in the experimental area ($N = 3000$). Standard deviations are in parentheses.