NBER WORKING PAPER SERIES

REFORMING INEFFICIENT ENERGY PRICING:
EVIDENCE FROM CHINA

Koichiro Ito
Shuang Zhang

Working Paper 26853
http://www.nber.org/papers/w26853

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2020

For helpful comments, we thank Hunt Allcott, Douglas Almond, Severin Borenstein, Fiona Burlig, Steve Cicala, Lucas Davis, Tatyana Deryugina, Larry Goulder, Michael Greenstone, Kelsey Jack, Ryan Kellogg, Catherine Wolfram, and seminar participants at Stanford, Colorado Boulder, UC Berkeley, University of Chicago, Resources for the Future, the Federal Reserve Board of Governors, and NBER Environment and Energy Program Meeting. We thank Theodor Kulczycki, Jing Qian, Chenyu Qiu, and Andrew Smith for excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Reforming Inefficient Energy Pricing: Evidence from China
Koichiro Ito and Shuang Zhang
NBER Working Paper No. 26853
March 2020
JEL No. L38,L51,L97,O1,O38,O44,Q4,Q41,Q5,Q53,Q56

ABSTRACT

Inefficient energy pricing hinders economic development in many countries. We examine long-run effects of a recent heating reform in China that replaced a commonly-used fixed-payment system with individually-metered pricing. Using staggered policy rollouts and administrative data on household-level daily heating consumption, we find that the reform induced long-run reductions in heating usage and generated substantial welfare gains. Consumers gradually learned how to conserve heating effectively, making short-run evaluations underestimate the policy impacts. Our results suggest that energy price reform is an effective way to improve allocative efficiency and air quality in developing countries, where unmetered-inefficient pricing is still ubiquitous.

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

Shuang Zhang
Department of Economics
University of Colorado Boulder
UCB 256
Boulder, CO 80302
shuang.zhang@colorado.edu

1 Introduction

Inefficient energy pricing is ubiquitous in developing countries, and it hinders progress on two fundamental factors for economic development—reliable energy access and clean air. Since energy prices are considered as powerful political instruments, energy consumption is often highly subsidized in developing countries, leaving marginal prices substantially lower than marginal costs. In 2018, global fossil-fuel consumption subsidies exceeded 400 billion dollars, and nearly all of them were used in developing countries (Information Energy Agency, 2020).[1] This inefficiency eventually results in limited and unreliable energy access for consumers in a dysfunctional equilibrium, where subsidies are high enough for firms to make profits without investing in infrastructure (McRae, 2015*b*). Furthermore, very low—and often zero—marginal prices provide little incentives to conserve energy, which exacerbates severe air pollution in many countries (Almond et al., 2009; Ebenstein et al., 2017; Ito and Zhang, 2020).

In this paper, we investigate long-run impacts of reforming inefficient energy pricing commonly used in developing countries. We collected newly-available administrative data on household-level daily heating usage for 16,133 households from 2007 to 2018 in the city of Tianjin, by collaborating with the World Bank, the Chinese government's Ministry of Housing and Urban-Rural Development (MOHURD), and a regulated heating provider. Residential heating bills in China have been solely based on annual fixed fees, regardless of each consumer's usage—that is, the marginal price has been subsidized to zero. Starting in 2005, the government installed individual meters and introduced new pricing called consumption-based billing (CBB): a two-part tariff with a constant marginal price and fixed fees.

Our research design and data address two key empirical challenges in the literature. First, individually-metered usage data are usually available *only after* the introduction of CBB because firms typically install meters at the same time as they introduce CBB (McRae, 2015*a*). It makes

[1] For example, the majority of households have paid zero marginal price for heating for a long time in China (Ito and Zhang, 2020). In India, liquefied petroleum gas price is highly subsidized (Barnwal, 2014), and a flat monthly fee with small or zero marginal price is common for electricity customers in the agricultural sector (Ryan and Sudarshan, 2018; Badiani-Magnusson and Jessoe, 2019). In Colombia, electricity pricing has been based solely on fixed charges until recently (McRae, 2015*a*). Turkmenistan is considering transition from a fixed payment scheme to consumption-based billing for residential electricity, natural gas and drinking water (Diplomat, 2018). Even in developed countries such as the United States, fixed payments with zero marginal price are still common for water and electricity in some regions or some customer segments, including apartment tenants and agricultural customers (Davis, 2012; Burlig, Preonas and Woerman, 2020).

empirical analysis challenging because individual-level usage data are unobservable before CBB. Our data overcome this challenge because regulators required household-level metered data to be collected at least one year before the introduction of CBB. This allows us to access daily household-level usage data *before and after* the reform. The second challenge in the literature is the lack of control group for estimating a causal effect of CBB. To address this problem, we develop an event-study design that exploits staggered rollouts of the reform.

We begin with graphical and statistical analysis of the event-study estimation. Our results on the intention-to-treat (ITT) indicate that the introduction of CBB reduced the average heating usage by 13% in the first year, 21% in the second year, 27% in the third year, and 31% in the fourth year, suggesting that heating demand became more elastic in the long run. In the fourth year, the average treatment effect on the treated (ATET) was a 36.3% reduction in heating usage. These impacts are economically substantial and long-lasting compared to a variety of policies on residential energy usage studied in the literature, as we discuss below.

We investigate two dimensions of the long-run dynamics in energy demand. First, we find that consumers were likely to gradually learn when they should conserve heating. We divide treatment days into quartiles of daily average temperature. The ATET is similar between colder and warmer days in the first year of treatment. However, after the second year of treatment, the response becomes larger on warmer days relative to colder days. It suggests that consumers gradually learn how to wisely save energy. Second, we find that relatively-poorer households became responsive to prices over time, while richer households responded similarly between the short run and long run.

To examine the welfare impacts of CBB, we need to test a hypothesis that has not been formally tested in the literature. A standard welfare analysis of two-part tariffs relies on the assumption that consumers properly distinguish fixed costs from variable costs. However, Liebman and Zeckhauser (2004) suggest that behavioral consumers may misperceive a change in average price as a change in marginal price in a complex price schedule. This behavior, if found in data, makes the standard welfare analysis for two-part tariffs no more appropriate. The price variation in our research design is suitable to test this hypothesis because many consumers experienced an *increase* in marginal price but a *decrease* in average price. We use this price variation to run the encompassing test (Davidson and MacKinnon, 1993; Ito, 2014). Our results suggest that consumers in our data properly distinguished fixed costs from variable costs, providing supporting evidence for an assumption in the standard welfare analysis of two-part tariffs.

Based on these empirical findings, we calculate welfare gains from CBB. We show that the social welfare gain from CBB comes from two elements: improved allocative efficiency and reduced environmental externalities. We separately estimate each welfare gain by integrating our empirical findings, plant-level emissions data, and willingness to pay for clean air estimated by Ito and Zhang (2020). Our results indicate that the total social welfare gain is 61.29 dollars per year per household and 262.68 million dollars per year for Tianjin. The one-time administrative cost of introducing CBB was 99.22 dollars per household, which implies that the policy's benefit exceeded its cost half way through the second treatment year.

In addition to the efficiency effects of CBB, another important welfare question is the redistributional impacts. To investigate this question, we collected data on housing prices as a proxy for wealth. We then calculate CBB's impacts on consumer surplus for wealthier households and poorer households separately. Our results suggest that the introduction of CBB increased consumer surplus on average, but it was regressive—the change in consumer surplus was higher for wealthier households than poorer households. This regressivity comes from the fact that the pre-reform policy was quite progressive. Because the fixed charge was proportional to the square meters of residences, the pre-reform payment was higher for wealthier households, who tended to live in larger residence.

Related literature and our contributions. First, our study provides the first evidence on the long-run responses to energy prices in developing countries. In the coming decades, most of the global increase in energy demand will come from developing countries (Wolfram, Shelef and Gertler, 2012). How to design energy pricing in these countries is, therefore, a first-order question for addressing climate change and global scarcity in natural resources. However, the energy demand literature has focused on developed nations because of the availability of administrative billing data.[2] Moreover, nearly all existing studies focus on estimating short-run demand elasticity because long-run exogenous variation in energy prices is hardly available.[3] We collected administrative

[2]For example, see Borenstein (2010), Aroonruengsawat and Auffhammer (2011), Wolak (2011), Ito (2014), Jessoe and Rapson (2014), Ito (2015), Ito, Ida and Tanaka (2018), and Deryugina, MacKay and Reif (2020) for studies based on administrative energy billing data in the United States and Japan. Recently, researchers started to collect such data in developing countries: Mexico (Davis, Fuchs and Gertler, 2014), South Africa (Jack and Smith, 2015, 2020), Colombia (McRae, 2015*a*), Brazil (Costa and Gerard, 2018), and Kenya (Lee, Miguel and Wolfram, 2020).

[3]Deryugina, MacKay and Reif (2020) emphasize this point and estimate two-year responses to electricity prices in Illinois. They find that Illinois households gradually responded to changes in electricity prices, which is consistent with our findings for Chinese households. Another related study is Costa and Gerard (2018), but their study focuses on persistent responses to a temporal policy shock, and therefore, different from Deryugina, MacKay and Reif (2020)

billing data in China, and our quasi-experimental design allows us to estimate four-year responses to long-run price variation. Our findings on the sizable and long-lasting welfare gains highlight that inefficiency in energy pricing is likely to be large in developing countries, and therefore, it is important to conduct rigorous studies in these countries.[4]

Second, our findings provide a new insight to the literature on consumer inattention under complex pricing. Motivated by Liebman and Zeckhauser (2004), many studies test if standard economic theory's assumptions are consistent with data when consumers have complex price schedules. For example, Ito (2014) finds that consumers respond to average price rather than marginal price under a nonlinear electricity price schedule, where the marginal price is a step function of monthly consumption, similar to multi-tiered income tax schedules. This evidence is, however often cited incorrectly to argue that consumers may be confused between marginal and average prices when a price schedule includes fixed and variable charges. Borenstein and Davis (2012) clarify that the evidence from Ito (2010, 2014) is not directly applicable to this argument, and what has to be tested is whether consumers distinguish fixed costs from variable costs. Our paper is the first study to test this hypothesis and find that consumers in our data are attentive to the difference between changes in fixed costs and variable costs.

Finally, our results provide important policy implications to prevailing severe air pollution in developing countries. Recent studies show that such severe air pollution causes large negative impacts on various economic outcomes, including infant mortality (Jayachandran, 2009; Arceo, Hanna and Oliva, 2012; Greenstone and Hanna, 2014), life expectancy (Chen et al., 2013; Ebenstein et al., 2017), and labor supply (Hanna and Oliva, 2015). For this reason, policymakers and economists consider air pollution to be one of the first-order obstacles to economic development. An important next step in this literature is to seek potential solutions to address this problem. Our paper is among the first studies to explore a solution in collaboration with policymakers. Our results highlight that reforming inefficient energy pricing is likely to be an effective way to alleviate air pollution in developing nations.

---

and our study.

[4]For example, Wolak (2011) and Ito, Ida and Tanaka (2018) find that the introduction of residential dynamic electricity pricing in the United States and Japan, which created increases in peak-hour prices by 100–300%, induced 10–15% reductions in electricity usage. Another policy that has been extensively studied in many developed countries is information provision with peer comparison in energy usage, which typically induced 1–2% reductions in energy use (Allcott and Rogers, 2014).

2 Research Design and Data

This section describes three key features of our research design. First, the heating pricing reform in Tianjin was introduced with staggered rollouts, which created quasi-experimental variation in treatment. Second, the policy-induced price variation allows us to test if consumers properly distinguished between fixed and variable costs. Third, the city required household-level metered usage data to be collected at least one year before the introduction of CBB, allowing us to access daily household-level usage data *before and after* the reform.

2.1 Staggered Rollouts

The city introduced CBB with staggered rollouts between 2008 and 2016. The timing of the rollout was well spread across years. The number of households who were introduced to CBB was 720 in 2008, 1,183 in 2009, 3136 in 2010, 1064 in 2011, 1209 in 2012, 2129 in 2013, 5531 in 2014, 738 in 2015, and 423 in 2016. This long window of rollouts enables us to estimate the policy's long-run effect based on event-study design.

The vast majority of households in Tianjin live in condominiums, and therefore, the rollouts were done at the condominium building level. Thus, in our estimation, we cluster the standard errors at the building level. The city's annual operation budget for the reform was constrained, which was the reason for the staggered rollouts spanned for nine years. According to the city officials, the rollouts were done in an unsystematic order, though the timing was not randomly assigned. Hence, our estimation relies on the standard identification assumptions for quasi-experimental event study design, as we describe in Section 3. For example, we test if the rollout timings are correlated with building characteristics. We do not find statistically significant relationships between the rollout timings and the observable building characteristics, including year of building, square meters of residence, and house value, as we discuss in Section 3.3.

Once a building was assigned to CBB, all households were defaulted into CBB. However, there was an option for them to opt-out from CBB and stay with the pre-reform tariff. They had to opt-out before the first winter of CBB. In our data, we observe that about 70 percent of households complied with CBB and 30 percent of households opted-out. For both compliers and non-compliers, we observe their daily metered heating usage. We address this incomplete compliance in our estimation in Section 3.

2.2 Price Variation Created by the Reform

Before the policy change, households paid an annual fixed charge equal to 3.97 dollars times their residence's square meters. For example, a household with 100 square meters of residence paid 397 dollars for every winter, regardless of how much heating was used.

After the policy change, a heating bill became a two-part tariff: a new annual fixed charge equal to 1.895 dollars times square meters; and a variable charge equal to 1.4 cents per kWh of heating usage.[5] This policy change provides useful variation for our empirical analysis because many consumers experienced an increase in marginal price but a decrease in average price. For example, consider a household with 100 square meters of residence whose typical usage was 10,000 kWh per winter. The pre-reform payment was 397 dollars with zero marginal price. With the same amount of usage, the post-reform payment would be 338.5 ($= 198.5 + 0.014 \cdot 10,000$) dollars with marginal price equal to 1.4 cents. Thus, given the same amount of usage, this household experienced an increase in marginal price but a decrease in average price.

The change in marginal price was common to all households—it changed from 0 to 1.4 cents per kWh. However, the change in average price depended on usage per square meter. Given the same usage level, households whose usage per square meter was less than 142 were likely to experience a decrease in average price and the rest of households were likely to experience an increase in average price. Based on the pre-reform usage data, we observe that 55% of households had usage per square meter less than 142. This implies that many consumers were likely to experience an *increase* in marginal price but a *decrease* in average price, which is key variation we use in Section 4.[6]

2.3 Data

We obtained administrative data on household-level daily heating usage from a regulated heating provider in Tianjin. The data cover usage for all of the provider's residential customers from December 2007 to February 2018. Heating usage is automatically recorded once a day and uploaded to the provider's database. With a confidentiality agreement, we obtained direct access to the

[5] The regulator set the marginal price equal to the marginal cost based on the information on heating production.
[6] Figure A.1 provides visual illustration of this price variation.

database. We also observe each household's address, apartment number, square meters, and house value.

The data include 16,133 households in total. However, as we explain in Section 3, our main analysis examines the long-run effects of CBB by analyzing households whose data are available at least one year before and four years after the introduction of CBB. With this restriction, our main dataset includes 3,874 households. In the appendix, we also show results on the short-run effects of CBB based on a dataset without this restriction and find that the short-run effects are similar between our main dataset and the dataset without the four-year data restriction (Figure A.4 and Table A.5).

Every year, a winter heating season starts in mid-November and ends in mid-March. Each year's exact start date in November and end date in March depend on that year's temperature. To make our analysis consistent between years, we focus on daily usage in three winter months—December, January, and February. In these three months, heating is on everyday in every year so that we do not have missing days for any year.

We report summary statistics for key variables here and other variables in Table A.1. The average heating usage is 99.92 kWh per day and 12,097.1 kWh per winter. The average heating bill per winter is 454.1 dollars before the reform and 396.7 dollars after the reform. The average size of residence is 114 square meters, and the average housing price is 539,682 dollars.

3 Treatment Effects of consumption-based billing

This section investigates the causal effect of consumption-based billing on heating usage. As we described in Section 2.1, the policy take-up was not mandatory, and about 70% of households complied. This is one-sided incomplete compliance because all households in the control group were untreated. For this reason, we estimate the intention-to-treat (ITT) and the average treatment effect on the treated (ATET).

3.1 Intention-to-Treat

We begin with a standard event study analysis, which provides visual investigation of treatment effects in the presence of staggered treatment assignment. We estimate the following equation by

OLS:

$$y_{it} = \alpha_i + \gamma_{st} + \sum_{k=a}^{b} \phi_k D_{it}^k + u_{it}, \tag{1}$$

where y_{it} is the natural log of daily heating usage for household i on day t, and α_i is household-level fixed effects (α_i). We allow day fixed effects (γ_{st}) to be different among households with different years of meter installations (s). We use $j = [a, b]$ to denote the *event-time* relative to the first month of treatment. For example, $k = -1$ is the last month of the pre-treatment period and $k = 0$ is the first month of treatment. Because we use data from three winter months—between the first day of December and the last day of February—it is helpful to consider the following example. Suppose a household whose first treatment month was December 2010. For this household, k equals -1 in February 2009, 0 in December 2010, 1 in January 2011, 2 in February 2011, 3 in December 2012, and etc. A dummy variable $D_{it}^k = 1$ if day t falls within the event-time k for household i.

We follow McCrary (2007) and Kline (2012) to address two important issues in event-study regression. First, we normalize $\phi_{-1} = 0$ so that all other ϕ_k are treatment effects relative to the the last month of the pre-treatment period. Second, we impose the following endpoint restrictions: $\phi_k = \underline{\phi}$ for $k < -3$ and $\phi_k = \overline{\phi}$ for $k > 11$ and make sure that all households in the analysis have balanced panel in the event-time within $k = [-3, 11]$. In this way, we make sure that ϕ_k in $k = [-3, 11]$ are estimated from the same set of households and comparable. Finally, u_{it} is the error term, and we cluster the standard error at the building level to account both for within-building correlation and serial correlation.

In Figure 1, we show the estimates of ϕ_k in $k = [-3, 11]$—between one year before and four years after the policy assignment. This figure provides three key results. First, there is no statistically significant pre-trend before the policy introduction. Second, the policy impact becomes larger in the second and third years after the policy implementation, and it persists in the fourth year. In Table 1, we estimate ϕ_k for each treatment year. The results indicates that the ITT was a reduction in heating usage by 13% in the first year, 21% in the second year, 27% in the third year, and 31% in the fourth year.[7]

[7] y_{it} is heating usage in log. Therefore, ϕ_k (-0.136, -0.230, -309, -0.367 in Table 1) provide the ITT in log points. The ITT in percentage changes can be obtained by $\exp(\phi_k) - 1$.

3.2 Average Treatment Effect on the Treated and the Heterogeneity

Under the standard assumptions for the local average treatment effect (LATE), the instrumental variable (IV) estimator provides the average treatment effect on the treated (ATET) because our incomplete compliance was only on the treatment group. To estimate the ATET, we include T_{it}^k (household i's actual treatment status in event-time k) as right-hand-side variables in equation (1), and use treatment assignment (D_{it}^k) as instruments. Column 2 in Table 1 suggests that the ATET was a reduction in usage by 18% in the first year, 28% in the second year, 33.4% in the third year, and 36.3% in the fourth year.[8]

These findings suggest evidence of learning—heating demand became more elastic in the long run. We investigate two dimensions of this long-run dynamics in energy demand by estimating heterogeneity in the ATET. In Panel A of Figure 2, we divide treatment days into quartiles of daily average temperature. The ATET is similar between colder and warmer days in the first year of treatment. However, after the second year of treatment, the response becomes larger on warmer days relative to colder days. It suggests that consumers gradually learn how to wisely save energy. In the absence of marginal incentives, they did not have to save heating in both warm and cold days. With a price incentive, it is reasonable to save more heating on warm days than cold days. The figure suggests that this behavior gradually occurred over time. In Panel B, we divide households into two groups by housing price, which is a proxy for wealth. we find that relatively-poorer households became responsive to prices over time, while richer households responded similarly between the short run and long run. This result indicates that learning is particularly sizable for less-wealthy households.

3.3 Validity of the Identification Assumptions.

The validity of our quasi-experiment is subject to a standard set of identification assumptions for event-study research design. The most important assumption in our estimation is the parallel trend in counterfactual outcomes. In the absence of treatment, the trajectory of the outcome variable (heating usage) has to be parallel between treatment and control groups. Although this is an untestable assumption, we provide two pieces of supporting evidence. The first evidence is the

[8] Because y_{it} is heating usage in log, ψ_k (-0.20, -0.33, -0.40, -0.45 in Table 1) are the ATET in log points. The ATET in percentage changes can be obtained by $\exp(\psi_k - 1)$.

absence of the pre-trend in Figure 1—the trajectory of heating usage was not statistically different between treatment and control groups prior to the treatment periods. Second, we test if observable building characteristics are associated with the the staggered timings of policy implementation. We do not find statistically significant relationships between the timings of policy implementation and the year of building, size of residence, and house value in Table A.2.

To interpret our IV estimate as the ATET, we also need the standard assumptions for the LATE (Imbens and Angrist, 1994). A potential concern is that the Stable Unit Treatment Value Assumption (SUTVA) can be violated if a household's usage is affected by other households' compliance decisions. We examine if we observe evidence of such violation in our data. We test if a change in household's heating usage is correlated with compliance decisions of neighbors, and do not find statistically significant correlation in Table A.4.

4 Demand Estimation and Encompassing Test

In this section, we use the policy-induced price variation to estimate demand for heating and implement the encompassing test (Davidson and MacKinnon, 1993; Ito, 2014) to examine if consumers in our data properly distinguished fixed costs from variable costs. Testing this hypothesis is key to the welfare analysis in Section 5.

4.1 Conceptual Framework

Before we present our empirical strategy, it is useful to provide a brief conceptual framework. Consider a utility maximization problem for heating demand y. A consumer has income I, the marginal price of heating is mp, and the fixed charge is f. We consider a quasi-linear utility function $u = v(y) - mp \cdot y - f + I$.[9]

A standard utility maximization problem simply solves the first order condition for the utility function, which leads to a condition: $v'(y^*) = mp$. Therefore, the optimal usage (y^*) can be obtained when the marginal utility from heating usage equals to the marginal price. On the other

[9] A quasi-utility function assumes that there is no income effect. This assumption is likely to be a valid in our empirical context because the income effect of the CBB policy was likely to be very small. The CBB policy reduced the annual fixed charge by about \$226 per household. The average household income in Tianjin in our sample period was \$15,041. Therefore, \$200 was about 1.5 percent of household income. In the literature on residential energy demand, the income elasticity is found quite small, around 0.01. It implies that the income effect of the CBB policy on usage would be a change in usage by about 0.015 percent.

hand, Liebman and Zeckhauser (2004) suggest the possibility that the consumer may misperceive a change in average price (ap) as a change in marginal price. Two-part tariffs could create this misperception if consumers do not properly distinguish between changes in fixed costs and changes in variable costs. In this case, the optimal usage can be characterized by $v'(y^{**}) = ap$.

In Section 2.2, we described that the policy introduction created different variation in changes in marginal and average prices (Figure A.1). We exploit this policy-induced variation to test whether consumers respond to marginal price or average price under a two-part tariff.

4.2 Empirical Analysis and Results of Encompassing Tests

Consider the following equation:

$$y_{it} = \alpha_i + \gamma_{st} + \beta_1 MP_{it} + \beta_2 AP_{it} + u_{it}, \tag{2}$$

where the dependent variable (the log of heating usage) and fixed effects are the same as equation (1). MP_{it} and AP_{it} are the marginal and average prices of heating usage. In the encompassing test, we examine how the inclusion of one variable or another (MP_{it} and AP_{it}) affects the estimates of β_1 and β_2. If consumer behavior is more consistent with standard economic theory, the inclusion of AP_{it} would not affect the estimate of β_1, and the estimate of β_2 would be near zero.[10] We expect an opposite result if consumer behavior is more consistent with shmeduling (Liebman and Zeckhauser, 2004).

The price variables are endogenous for two reasons. First, the policy's incomplete compliance implies that self-selection into the new prices can be correlated with u_{it}. Second, AP_{it} is a function of usage, which creates a simultaneity between usage and prices (Borenstein, 2009; Ito, 2014).

To address this problem, we use *predicted* prices based on policy-induced variation (MP_{it}^{PI} and AP_{it}^{PI}) as instruments. MP_{it}^{PI} takes zero for the pre-policy period and 0.014 USD for the post-policy period, regardless of customer i's compliance. In this way, the instrument is not subject to endogeneity from incomplete compliance. As discussed in Section 2.2, the policy-induced variation in average price is substantially different between consumers with lower usage per square meter and those with higher usage per square meter. To capture this variation, we construct AP_{it}^{PI}, following

[10] This may not be the case if the income effect is large enough, although this is unlikely in our case as we described in footnote 9.

the literature on nonlinear income taxation and nonlinear pricing (Saez, Slemrod and Giertz, 2012; Ito, 2014). We use $\widetilde{y}_i$ to denote customer i's average daily usage per square meter prior to the first treatment year. We calculate the policy-induced predicted average price AP_{it}^{PI} with household i's $\widetilde{y}_i$ with the price schedule at time t.

The first stage of this instrument is strong because a consumer's past usage is a strong predictor for their future usage (Chetty et al., 2011). However, Saez, Slemrod and Giertz (2012) and Ito (2014) emphasize that the exclusion restriction is violated if the mean reversion of the outcome variable is not properly controlled. In general, economic data—such as energy usage or income earnings—show strong mean reversion. Customers with low $\widetilde{y}_i$ are likely to increase their usage in other periods, while those with high $\widetilde{y}_i$ are likely to decrease their usage in other periods, for reasons unrelated to their responses to prices.

To address this concern, we take a nonparametric control approach, following Ito (2014). The idea is that if there is a valid control group, one can include flexible controls for the mean reversion, assuming that the mean reversion is not systematically different between treatment and control groups. We make percentile groups ($k = 1, ..., 100$) based on $\widetilde{y}_i$. Then, we allow the day fixed effects (γ_{st}) to differ by the percentile groups. This flexible control does not absorb the policy-induced price variation in our research design because the event-study design provides control and treatment groups within each percentile group.

Columns 3 of Table 1 shows that the semi-elasticity with respect to marginal price is -14.9. Because the dependent variable is the log of heating usage and the price variables are in dollars, the estimate implies that one cent increase in marginal price would result in a 14.9 percent reduction in usage. In column 4, we include both marginal and average prices. Including average price merely changes the coefficient for marginal price, suggesting that the variation in average price does not explain much about the variation in usage conditional on marginal price. This result suggests that data are consistent with standard economic theory—consumers properly distinguish between fixed and variable costs—and supports that a standard welfare analysis can be applied in Section 5.

5 Welfare Analysis

In Figure 3, we explain how consumption-based billing affects social welfare and consumer surplus. Before the policy change, the regulated heating provider offered a tariff with zero marginal price and a fixed charge. Then, the regulated provider introduced CBB as a form of two-part tariff with a marginal price that was set equal to the marginal cost. In the figure, Y_0 and Y_1 are usage before and after the introduction of CBB. A social welfare gain comes from the improved allocative efficiency (C in the figure). In addition, suppose that there are unpriced environmental externalities. Then, CBB creates another welfare gain (D in the figure) from reduced externalities.

Consumer surplus can be increased or decreased by CBB. The annual fixed charges were $3.97 \cdot s$ dollars before the reform, where s is square meters of residence, and $\frac{1}{2} \cdot 3.97 \cdot s$ dollars after the reform. Hence, for a quasi-linear utility function, the change in consumer surplus equals to $\frac{1}{2} \cdot 3.97 \cdot s - (A + B)$ in the figure. In this section, we calculate social welfare gain and changes in consumer surplus based on our empirical findings in the previous sections.

5.1 Social Welfare

In Panel A of Table 2, we calculate social welfare gains based on the ITT estimates in Table 1. The allocative efficiency (C in Figure 3) is 30.23 dollars per household per year and 129.54 million dollars per year for Tianjin.

To calculate an additional welfare gain from reduced environmental externalities (D in Figure 3), we need to know the externalities from heating production. We make the following steps to provide an estimate. First, we leverage the heating provider's plant-level hourly emissions data from the nation-wide continuous emission monitoring systems (CEMS) to estimate the relationship between heating consumption and emissions. As we report in Table A.3, we find that 1 percent increase in heating consumption is associated with 0.523 percent increase in PM_{10} concentrations, where the baseline PM_{10} concentrations before the reform was 143. This implies that our ITT estimate (31 percent reductions in usage) induced a reductions in PM_{10} concentrations by 23.185 ug/m^3. Second, Ito and Zhang (2020) suggests that a Chinese household's marginal willingness to pay for a reduction in PM_{10} is 1.34 dollars per ug/m^3 of PM_{10} per year. Then, multiplying these two estimates imply that the WTP for the policy-induced reduction in PM_{10} is 31.07 dollars per household per year and 133.15 million dollars per year for Tianjin. Note that this estimate is likely to be a lower bound estimate for environmental externalities because this calculation does not include other potential environmental externalities than PM_{10} and the MWTP for reductions in PM_{10} in Ito and Zhang (forthcoming) is a lower bound estimate for reasons described in that study.

The total social welfare gain ($C + D$ in Figure 3) is 61.29 dollars per household per year and 262.68 million dollars per year for Tianjin. To put this number in context, we also show that this social welfare gain is 16.6 percent of the heating provider's pre-reform total annual revenue. The one-time administrative cost of introducing CBB including the cost of installing meters was 99.22 dollars per household, which implies the policy's benefit exceeded its cost halfway through the second treatment year.

5.2 Consumer Surplus

In Panel B of Table 2, we divide households into two groups by their housing prices, and calculate the changes in consumer surplus for each group. Both groups had increases in consumer surplus, and the increase was larger for households with higher housing prices. Similarly, the payment for heating decreased for both groups, with a larger decrease for households with higher housing prices.

These results provide two key implications. First, the policy produced overall gains in consumer surplus in addition to the social welfare gains. This is a policy-relevant finding because how the policy's impacts on consumers matter to the adaptation of consumption-based billing. Second, the results suggest that the introduction of consumption-based billing seems regressive in the sense that gains in consumer surplus were larger for the rich. However, a more accurate statement is that the pre-existing policy—a fixed charge based on square feet of residence—was quite progressive because it was essentially a cross-subsidy from households with larger homes to those with smaller homes. The introduction of consumption-based billing lowered this cross-subsidy, and therefore, seemed more regressive than the pre-existing policy.

6 Conclusion

In this paper, we examined long-run effects of a recent heating reform in China that replaced a commonly-used fixed-payment system with individually-metered pricing. Using staggered policy rollouts and administrative data on household-level daily heating consumption, we found that the reform induced long-run reductions in heating usage and generated substantial welfare gains. Consumers gradually learned how to conserve heating effectively, making short-run evaluations underestimate the policy impacts. Our results suggest that energy price reform is an effective way to improve allocative efficiency and air quality in developing countries, where unmetered-inefficient pricing is still ubiquitous.

References

Allcott, Hunt, and Todd Rogers. 2014. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." *American Economic Review*, 104(10): 3003–37.

Almond, Douglas, Yuyu Chen, Michael Greenstone, and Hongbin Li. 2009. "Winter Heating or Clean Air? Unintended Impacts of China's Huai River Policy." *The American Economic Review: Papers and Proceedings*, 99(2): 184–190.

Arceo, Eva, Rema Hanna, and Paulina Oliva. 2012. "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City." NBER Working Paper No. 18349.

Aroonruengsawat, Anin, and Maximilian Auffhammer. 2011. "Impacts of Climate Change on Residential Electricity Consumption: Evidence from Billing Data." National Bureau of Economic Research, Inc NBER Chapters.

Badiani-Magnusson, Reena, and Katrina Jessoe. 2019. "Electricity Prices, Groundwater, and Agriculture: The Environmental and Agricultural Impacts of Electricity Subsidies in India." , ed. Wolfram Schlenker, Chapter Agricultural Productivity and Producer Behavior, 157–183. University of Chicago Press.

Barnwal, Prabhat. 2014. "Curbing leakage in public programs with direct benefit transfers." *Working Paper.*

Borenstein, Severin. 2009. "To What Electricity Price Do Consumers Respond? Residential Demand Elasticity Under Increasing-Block Pricing." *Working Paper.*

Borenstein, Severin. 2010. "The Redistributional Impact of Non-linear Electricity Pricing." *National Bureau of Economic Research Working Paper*, No. 15822.

Borenstein, Severin, and Lucas Davis. 2012. "The equity and efficiency of two-part tariffs in US natural gas markets." *The Journal of Law and Economics*, 55(1): 75–128.

Burlig, Fiona, Louis Preonas, and Matt Woerman. 2020. "Spatial externalities in groundwater extraction: Evidence from California agriculture." *Working Paper.*

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. 2013. "Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy." *Proceedings of the National Academy of Sciences*, 110(32): 12936–12941.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *The Quarterly Journal of Economics*, 126(2): 749–804.

Costa, Francisco, and François Gerard. 2018. "Hysteresis and the welfare effect of corrective policies: Theory and evidence from an energy-saving program." National Bureau of Economic Research.

Davidson, Russell, and James G. MacKinnon. 1993. "Estimation and Inference in Econometrics." Oxford University Press OUP Catalogue.

Davis, Lucas W. 2012. "Evaluating the Slow Adoption of Energy Efficient Investments Are Renters Less Likely to Have Energy Efficient Appliances?" , ed. Don Fullerton and Catherine Wolfram Vol. The Design and Implementation of U.S. Climate Policy. University of Chicago Press.

Davis, Lucas W., Alan Fuchs, and Paul J. Gertler. 2014. "Cash for Coolers: Evaluating a Large-Scale Appliance Replacement Program in Mexico." *American Economic Journal: Economic Policy*, 6(4).

Deryugina, Tatyana, Alexander MacKay, and Julian Reif. 2020. "The long-run dynamics of electricity demand: Evidence from municipal aggregation." *American Economic Journal: Applied Economics*, 12(1): 86–114.

Diplomat. 2018. "Turkmenistan Set to Rollback Subsidies for Good."

Ebenstein, Avaraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. 2017. "New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy." *Proceedings of the National Academy of Sciences*, 114(39): 10384–10389.

Greenstone, Michael, and Rema Hanna. 2014. "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India." *American Economic Review*, 104(10): 3038–3072.

Hanna, Rema, and Paulina Oliva. 2015. "The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City." *Journal of Public Economics*, 122: 68–79.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, 62(2): 467–475.

Information Energy Agency. 2020. "Energy subsidies: Tracking the impact of fossil-fuel subsidies."

Ito, Koichiro. 2010. "How Do Consumers Respond to Nonlinear Pricing? Evidence from Household Electricity Demand." *Working Paper University of California, Berkeley, Energy Institute at Haas.*, , (210).

Ito, Koichiro. 2014. "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing." *American Economic Review*, 104(2): 537–63.

Ito, Koichiro. 2015. "Asymmetric Incentives in Subsidies: Evidence from a Large-Scale Electricity Rebate Program." *American Economic Journal: Economic Policy*, 7(3): 209–37.

Ito, Koichiro, and Shuang Zhang. 2020. "Willingness to Pay for Clean Air: Evidence from Air Purifier Markets in China." *Journal of Political Economy.*

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.

Jack, B. Kelsey, and Grant Smith. 2015. "Pay as you go: Prepaid metering and electricity expenditures in South Africa." *American Economic Review*, 105(5): 237–41.

Jack, B. Kelsey, and Grant Smith. 2020. "Charging ahead: Prepaid electricity metering in South Africa." *American Economic Journal: Applied Economics.*

Jayachandran, Seema. 2009. "Air Quality and Early-Life Mortality: Evidence from Indonesia's Wildfires." *Journal of Human Resources*, 44(4): 916–954.

Jessoe, Katrina, and David Rapson. 2014. "Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use." *American Economic Review*, 104(4): 1417–38.

Kline, Patrick. 2012. "The Impact of Juvenile Curfew Laws on Arrests of Youth and Adults." *American Law and Economics Review*, 14(1): 44–67.

Lee, Kenneth, Edward Miguel, and Catherine Wolfram. 2020. "Experimental Evidence on the Economics of Rural Electrification." *Journal of Political Economy.*

Liebman, Jeffrey, and Richard Zeckhauser. 2004. "Schmeduling." *www.hks.harvard.edu/jeffreyliebman/schmeduling.pdf.*

McCrary, Justin. 2007. "The Effect of Court-Ordered Hiring Quotas on the Composition and Quality of Police." *American Economic Review*, 97(1): 318–353.

McRae, Shaun. 2015*a*. "Efficiency and Equity Effects of Electricity Metering: Evidence from Colombia." *Working Paper.*

McRae, Shaun. 2015*b*. "Infrastructure quality and the subsidy trap." *American Economic Review*, 105(1): 35–66.

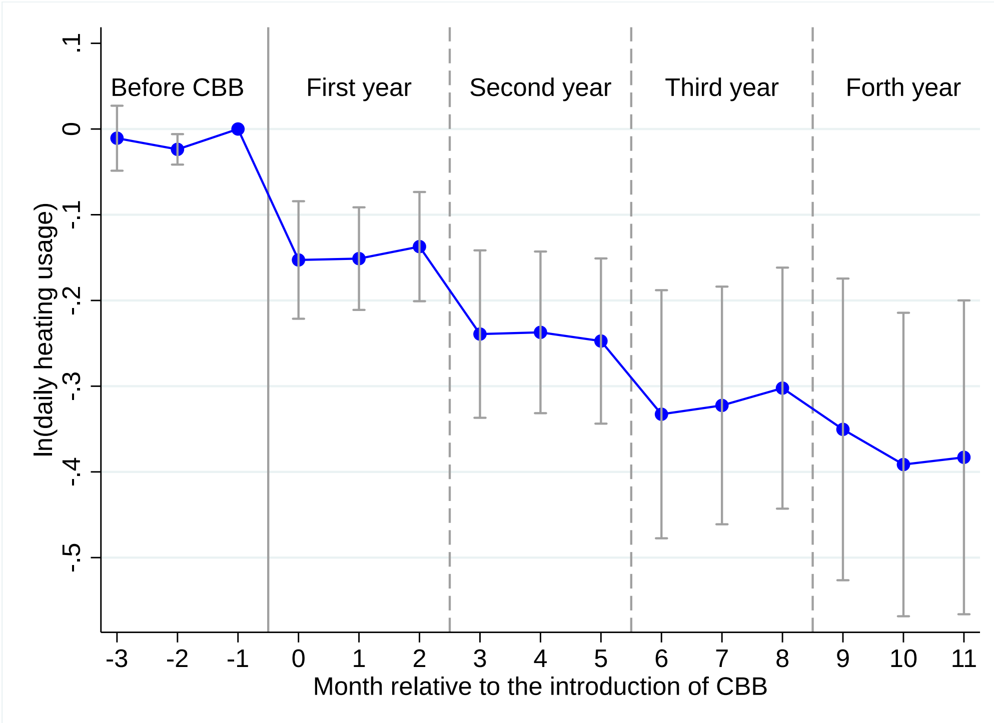
Ryan, Nicholas, and Anant Sudarshan. 2018. "The Efficiency of Rationing: Agricultural Power Subsidies, Power Supply and Groundwater Depletion in Rajasthan." *Working Paper.*

Saez, Emmanuel, Joel B. Slemrod, and Seth H. Giertz. 2012. "The elasticity of taxable income with respect to marginal tax rates: A critical review." *Journal of Economic Literature*, 50(1): 3–50.

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.

Wolfram, Catherine, Orie Shelef, and Paul Gertler. 2012. "How will energy demand develop in the developing world?" *Journal of Economic Perspectives*, 26(1): 119–38.

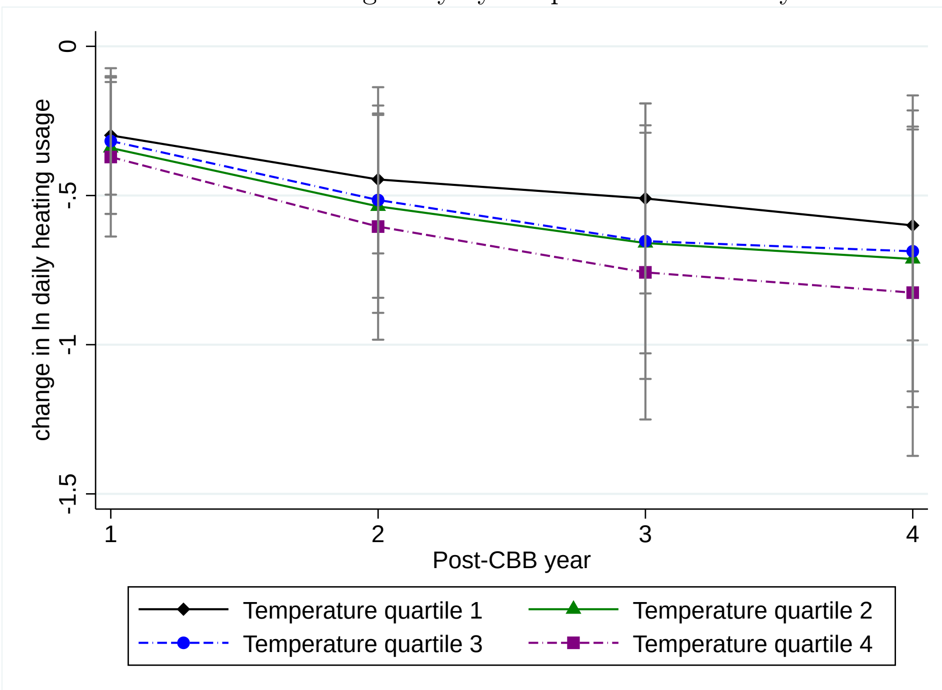
Figure 1: Event-Study Analysis: Intention-to-Treat (ITT)



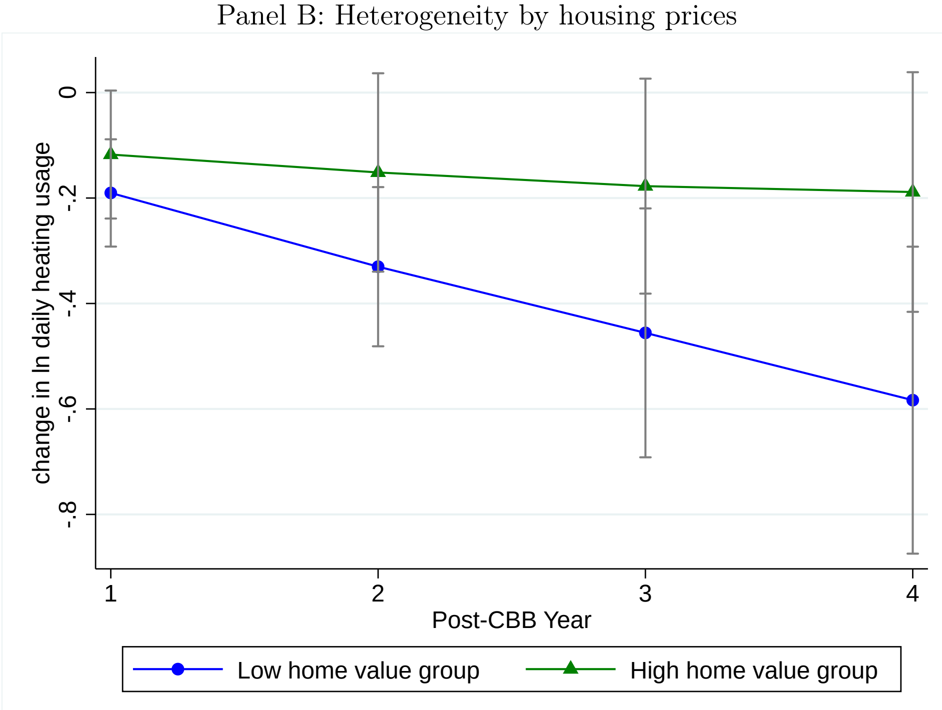
Notes: This figure shows the ITT estimates from equation (1). There are three months in each treatment year because the heating season is December, January, and February. The bars indicate the 95 percent confidence intervals. Standard errors are clustered at the building level.

Figure 2: Heterogeneity in the Average Treatment Effects on the Treated (ATET)

Panel A: Heterogeneity by temperature of the day

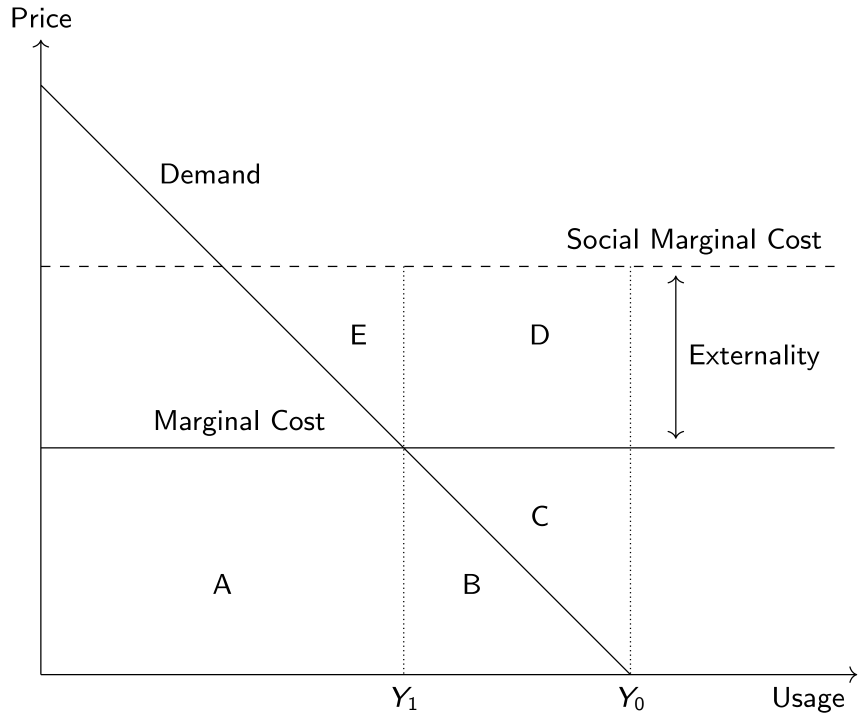


Panel B: Heterogeneity by housing prices



Notes: Panel A shows the ATET by quartiles of daily temperature. The thresholds for the quartiles are 25.8, 29.9, and 33.7 F. Panel B shows the ATET for households whose house values are lower than the median in our data (512,885 dollars) and those whose housing prices are higher than the median.

Figure 3: Welfare Impacts of Consumption-Based Billing



Notes: This figure shows the conceptual framework for the welfare impacts of consumption-based billing. Y_0 and Y_1 are usage before and after the introduction of CBB. A social welfare gain comes from the improved allocative efficiency (C in the figure) and reduced environmental externalities (D in the figure). A change in consumer surplus is the change in the fixed payment plus $-(A+B)$ in the figure.

Table 1: Main Estimation Results

	Dependent variable: Log of daily heating usage			
	(1)	(2)	(3)	(4)
	ITT	ATET	Encompassing Test	
First year of treatment	-0.136	-0.205		
	(0.029)	(0.044)		
Second year of treatment	-0.230	-0.331		
	(0.047)	(0.077)		
Third year of treatment	-0.309	-0.407		
	(0.069)	(0.109)		
Fourth year of treatment	-0.367	-0.451		
	(0.089)	(0.148)		
Marginal price			-14.916	-14.928
			(1.359)	(1.348)
Average price				-0.255
				(0.353)
Observations	2,332,745	2,332,745	2,332,745	2,332,745
R^2	0.66			
First-Stage F-Stat		17.12	2135.55	65.11
Estimation	OLS	IV	IV	IV
Household FE	Y	Y	Y	Y
Day*First data year FE	Y	Y		
Day*First data year FE*Pct FE			Y	Y

Notes: This table shows the estimation results from equations (1) and (2). Standard errors in parentheses are clustered at the building level.

Table 2: Welfare Gains from Consumption-Based Billing

Panel A: Social welfare		
	Welfare gain per household (USD/year)	Welfare gain for Tianjin (million USD/year)
Social welfare gain from improved allocative efficiency (C in Figure 3)	30.23	129.54
Social welfare gain from reduced environmental externalities (D in Figure 3)	31.07	133.15
Total social welfare gain (C+D in Figure 3)	61.29	262.68
Social welfare gain relative to pre-reform total revenue		16.6%

Panel B: Consumer surplus		
	Households with lower home value	Households with higher home value
Consumer surplus gains (USD/year)	0.25	26.72
Percentage changes in payment relative to pre-reform	-6%	-13%

Notes: See texts in section 5 how we calculate the welfare gains and consumer surplus.

Online Appendices: Not For Publication

A Additional Tables

Table A.1: Summary Statistics

Daily heating usage (kWh)	106.27
	(49.15)
Total heating usage per heating season (kWh)	13,092.5
	(4,760.7)
Heating bill per heating season before reform (dollar)	469.2
	(148.9)
Heating bill per heating season after reform (dollar)	419.1
	(129.5)
Square meter of the residence	118.2
	(37.51)
Housing price (dollar)	544,523
	(232,330)
Take-up rate of the CBB policy	0.64
	(0.48)
Number of households	3,874
Observations	2,332,754

Table A.2: Rollout Timing and Building Characteristics

Dependent variable: Rollout year of usage-based pricing	
Year of build	0.008
	(0.029)
Average condo size (square meter)	-0.001
	(0.002)
Average home price per square meter (1,000 dollars)	0.028
	(0.036)
First meter data year FE	Y
Observations	171
R^2	0.84

Notes: In this table, we test if observable building characteristics are associated with the the staggered rollout timings of policy implementation. The observations are at the building level. The dependent variable is the rollout year of consumption-based billing.

Table A.3: Household Heating Usage and Plant Emissions

	(1) lnSO2	(2) lnNOX	(3) lnPM
ln(Daily total heating usage)	1.032***	0.700***	0.523***
	(0.339)	(0.227)	(0.145)
Observations	270	270	270
R^2	0.72	0.94	0.82
Year-month FE	Y	Y	Y

Notes: In this table, we estimate the relationship between heating usage and plant-level emissions. We observe high frequency (hourly) emissions that come from the main heating plant in the heating provider. The data come from the nation-wide continuous emission monitoring systems (CEMS). From 2014, high-emitting firms were required to upload hourly, automatically recorded pollutant-specific concentration data to a publicly available, online platform for each province. Plant-level monitors are installed on stacks associated with generating units and perform hourly measurements of the concentration of SO_2, NO_x and PM in emitted stack gases. To examine the relationship between household heating usage and plant emissions, we obtain hourly emission data from the heating plant. For each of the three pollutants, we measure the plant's average daily concentration by the average of all observed hourly values in a 24-hour period. Figure A.3 shows the scatter plot of these three variables, controlling for year-month fixed effects. Daily total heating usage is positively correlated with plant-level daily SO_2 concentration, and similarly with NO_x and PM. In this table, we control for year-by-month fixed effects and find that, a 1% increase in daily total heating usage is associated with a 1.03% increase in the plant's daily SO_2 emission concentration, a 0.7% increase in NO_x concentration and a 0.52% increase in PM concentration. The emission concentrations are measured on stacks before the combustion gases go into the air, and therefore local weather conditions do not affect the measures directly.

Table A.4: Does Heating Usage Depends on Neighbors' Compliance Status?

	ln(daily heating usage)
CBB*complied	-0.089***
	(0.024)
CBB*next door neighbor complied	-0.008
	(0.020)
CBB*upper level neighbor complied	0.010
	(0.015)
CBB*lower level neighbor complied	0.003
	(0.015)
Observations	2,331,098
R^2	0.71
Day*Building FE	Y
Household FE	Y

Notes: In this table, we test if changes in heating usage are correlated with neighbors' compliance status. The dependent variable is log of daily heating usage at the household level. The first coefficient implies that changes in heating usage are negatively correlated with households own compliance status, which is consistent with our main findings on the policy's treatment effects. The rest of the coefficients indicate that there is little statistical evidence that changes in heating usage are correlated with neighbors' compliance status.

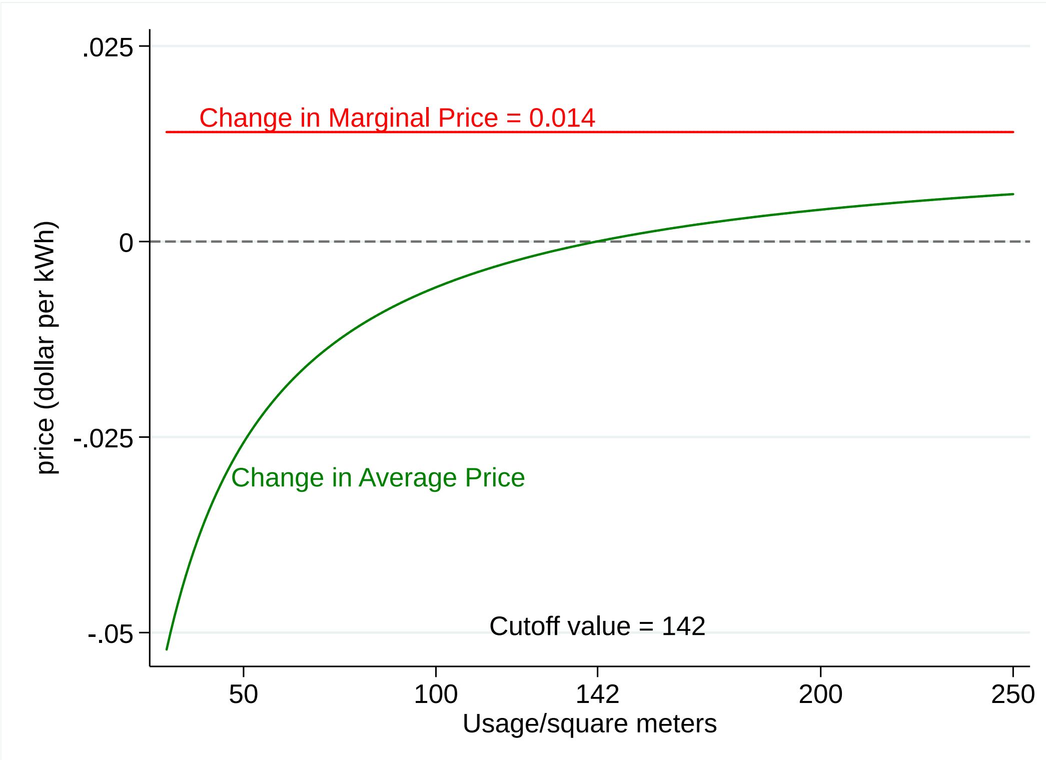
Table A.5: Main Estimation Results: Based on Households Whose Data Are Available One Year Before and One Year After the CBB

Dependent variable: Log of daily heating usage		
	(1) ITT	(2) ATET
First year of treatment	-0.115	-0.170
	(0.023)	(0.033)
Observations	3,258,304	3,258,304
R^2	0.66	0.66
First-Stage F-Stat		253.06
Estimation	OLS	IV
Household FE	Y	Y
Day*First data year FE	Y	Y

Notes: This table shows the estimation results from equations (1) and (2) using households whose data are available one year before and one year after the CBB. Standard errors in parentheses are clustered at the building level.

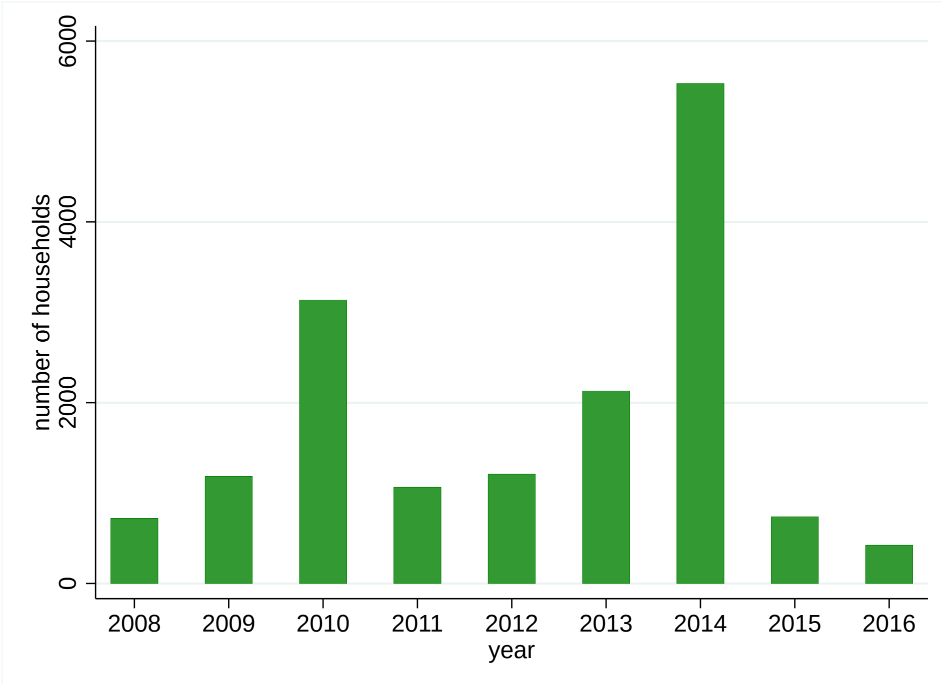
B Additional Figures

Figure A.1: Policy-Induced Changes in Marginal and Average Prices



Notes: This figure shows the changes in marginal price and average price induced by the introduction of the consumption-based billing policy.

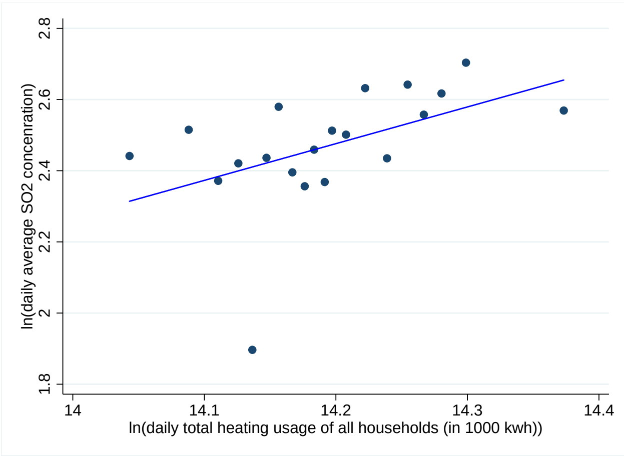
Figure A.2: Rollout Timings of Consumption-Based Billing



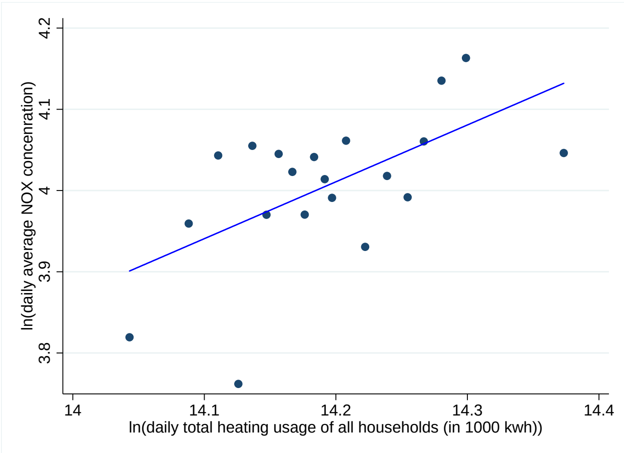
Notes: This figure shows the rollout of the consumption-based billing policy.

Figure A.3: Household Heating Usage and Plant Emissions

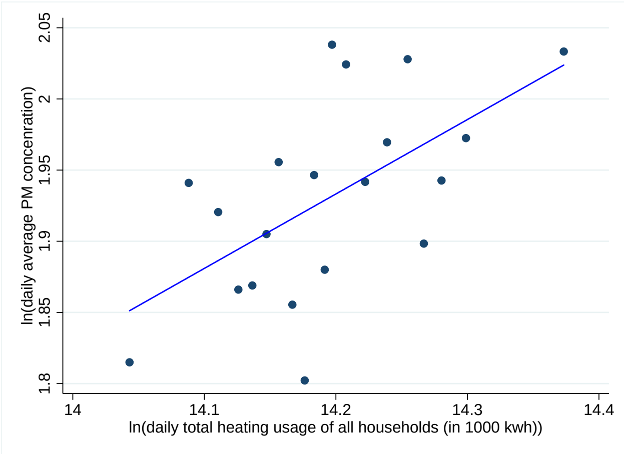
(a) SO_2 emission concentration (mg/m^3)



(b) NO_x emission concentration (mg/m^3)

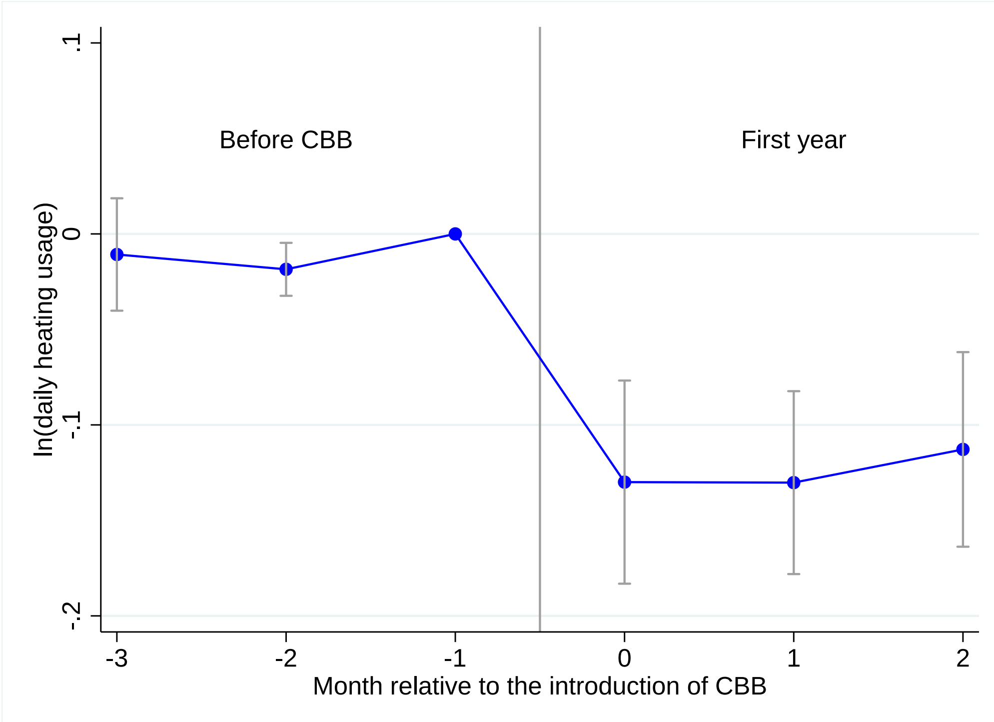


(c) PM emission concentration (mg/m^3)



Notes: See notes in Table A.3.

Figure A.4: Event-Study Analysis: Intention-to-Treat (ITT): Based on Households Whose Data Are Available One Year Before and One Year After the CBB



Notes: This figure shows the ITT estimates from equation (1) using households whose data are available one year before and one year after the CBB. The bars indicate the 95 percent confidence intervals. Standard errors are clustered at the building level.