Non-payment and strategic enforcement: Experimental evidence from Indian electricity

Pre-analysis plan

Fiona Burlig and Anant Sudarshan*

December 24, 2021

Many developing countries struggle to collect public revenue, reducing their ability to provide services. An important example of this is electricity distribution in India. Electricity is supplied to people largely by state-run utilities that regularly incur heavy losses because they struggle to collect on payments due to them. In turn, these bankrupt utilities resort to widespread power rationing to stay afloat. In partnership with two state utilities in Madhya Pradesh, India, we plan to run an experiment to measure the effectiveness of different strategies to enforce payments, implemented in a setting where approximately 65 percent of domestic consumers have outstanding arrears. Specifically, we randomize 30,000 utility consumers into a control group and multiple treatment arms: (i) SMS reminders (with and without peer comparisons), (ii) Warning notices sent by post, (iii) Warning notices delivered in person, and (iv) Disconnections for non-payment. We plan to compute direct effects of these treatments on bill payments and electricity consumption. We will also estimate treatment effect heterogeneity. Finally, we will conduct an endline survey to measure household finances, consumption, psychological well-being, and other measures.

This pre-analysis plan describes our experimental design, data collection, and outlines our planned analysis, including definitions of the key outcome variables and main regression specifications. We leave open the possibility of additional analysis informed by what we learn through implementation, and this plan should therefore not be viewed as comprehensive.

Key words: Enforcement; bill non-payment; electricity
JEL Codes: O13; Q48; H26

---

*Burlig: Harris School of Public Policy and Energy Policy Institute (EPIC), University of Chicago, and NBER. Email: burlig@uchicago.edu. Mailing address: Keller Center, 1307 E 60th St., Chicago, IL 60637. Sudarshan: Department of Economics and EPIC, University of Chicago. Email: anants@uchicago.edu. Mailing address: 5757 S. University Avenue, Chicago, IL 60637. We thank Susanna Berkouwer, Judd Boomhower, Michael Greenstone, Koichiro Ito, Erin Kelley, Ryan Kellogg, Gregory Lane, Catherine Wolfram, and seminar participants at Midwest Energy Fest and UC Energy Camp for helpful comments and suggestions. We gratefully acknowledge financial support from the Oak Foundation and the Griffin Applied Economics Incubator. Animesh Jayant, Simran Kalra, Garrison Schlauch, and Rathan Sudheer provided excellent research assistance. All remaining errors are our own. This project received IRB approval from the University of Chicago (Protocol No. IRB20-2127), and is registered on the AEA RCT registry (Identification No. AEARCTR-0008742). All remaining errors are our own.

Contents

1 Experimental design

In December 2021, we initiated a randomized controlled trial with the Madhya Pradesh Power Management Company Limited (MPPMCL) — the holding company for the five state-owned distribution utilities in the state of Madhya Pradesh in India. The goal of our collaboration is to test the effectiveness of different interventions designed to increase the payment rate of the utility consumer base. In what follows we describe different elements of the experiment in more detail.

1.1 Sample

Our focus in this project is on interventions that the utility could deploy at the level of an individual consumer. Accordingly the experimental sample consists of individual consumers. The first step in sample selection involved identifying two *circles*, Hoshangabad and Narsinghpur, as the site of the experiment. A *circle* is an operational division used by the utilities (discoms).[1]

We selected these two circles following the advice of MPPMCL, who identified them because they are administered by two different state-run utilities, had relatively good quality administrative data, represented a diverse consumer mix, and were not in violence-affected areas.[2]

Having identified our geography (circles) of interest, we began with a 10 percent random sample of the universe of registered *residential* consumers in the two circles. Residential (household) consumers were identified using the tariff classes associated with every consumer ID. Thus connections billed on agricultural, industrial, business, or government rates were excluded. After drawing this 10 percent sample we made three additional restrictions before selecting the experiment sample.

Drop unmetered and BPL customers We removed all households with an unmetered electricity connection, including households on a below-poverty-line (BPL) electricity tariff. We excluded these households for several reasons. These customers are charged a flat tariff and metering accuracy is therefore suspect since billing does not depend on consumption — meaning that a key outcome variable, electricity use, would be measured with substantial error for these customers. These unmetered households are also more likely to live in remote rural areas, making it operationally more expensive to disconnect them for persistent non-payment (one of our treatment arms, as we describe below). Finally, unmetered households, and especially BPL customers, consume less

1. In March 2021 — before the experiment was initiated — the circle of Hoshangabad was split into two distinct circles: Hoshangabad and Harda. The name Hoshangabad in this document refers to their union.

2. Madhya Pradesh suffers from violence linked to domestic insurgencies in certain areas which hampers governance on several dimensions and also would have made fieldwork and data collection dangerous.

energy on average on than metered households and contribute relatively little to revenue losses on an individual basis (both consumption and the monthly flat tariff faced by these consumers is low).

Drop inactive customers We removed all customers who did not consume electricity in August 2021. We had access to billing data for all customers for August 2021 while designing the sampling frame and we removed zero consumption records before drawing the sample to avoid including inactive customers.

Drop customers with no contact details Some of the treatment interventions and data collection require either a phone number or an address. The utility dataset we were provided contained contact details only for consumers who joined in early 2020 or before. By constraining ourselves to administrative IDs for whom contact details are available, we were forced to drop consumers with new connections.

Drop minor offenders On the direction of our government partners, we removed households whose August 2021 arrears were less than Rs 1,200. This ensured that no enforcement actions did not target customers who do not have substantial arrears. This restriction acknowledges the practical difficulties for the utility in targeting such consumers.

These sampling restrictions applied to the ten percent sample left us with 51,252 customers to draw from. We then randomly sampled 15,000 customers from each circle, for a total of 30,000 customers, to form the experimental sample.

1.2 Randomization and treatment arms

We randomly assigned each of these 30,000 customers to one of seven experimental arms, described below. We stratified our randomization on three factors.[3] First, we stratified on distribution centre (DC), the smallest administrative unit in the utility, and the local office level at which billing, collections, and disconnections are carried out. There are 57 DCs in our sample. Second, we stratified on whether a consumer uses more or less electricity than the median. Finally, we stratified on terciles of arrears. These final two strata are based on August 2021 data. This stratification approach yields 333 individual groups and increases statistical power and guides our approach to heterogeneous effects analysis, described below.

We next describe the experimental arms. Note that each of these treatments was selected from a larger menu of options provided by the researchers to MPPMCL. This design therefore excludes options that the government felt were undesirable or infeasible. For example, both increasing late fees and a one-time debt waiver were proposed and rejected.

3. We use the `-randtreat-` function in Stata to perform this stratified randomization.

A. Pure control: 7,500 consumers were randomized into the pure control group, where utility employees are requested to engage in no communication with households about payment beyond the details automatically provided on the electricity bill. This experimental arm approximates a no enforcement setting.

B. Business-as-usual: 2,500 consumers were randomized into a the business-as-usual (BAU) group, where utility employees are given no special requests. This is distinct from the pure control group, where the discom is told not to engage with customers.

C. SMS message: 5,000 consumers were randomized to receive repeated SMS messages over the 6 month implementation period reminding them that they are required to pay their bill. The text reads (translated to English from Hindi)

"Dear Consumer, please pay your electricity bill on time. There is [X] percentage of your bill remaining on your electricity connection number [IVRS No]. If payment is pending, please pay immediately to avoid disconnection."

D. SMS social comparison message: 5,000 consumers will receive an SMS message comparing their payment behavior to that of their neighbors.

"Dear Consumer, please pay your electricity bill on time. There is [X] percentage of your bill remaining on your electricity connection number [IVRS No]. Other similar consumers have [Y] percentage of their bill pending. If payment is pending, please pay immediately to avoid disconnection."

These messages (Arms C and D) will be sent by the utilities' IT department. These SMS reminders form a new addition to existing utility practice and were intended to test if lighter touch approaches to enforcement that do not involve physical disconnections could nevertheless be effective. For the purpose of calculating peer performance in message D, the definition of 'similar consumers' is other households in the same strata as the SMS recipient.

E. Mailed disconnection notice, no disconnections: 5,000 consumers will be mailed a disconnection notice, an official document used by the utilities in the course of normal business operations. The text reads (translated to English from Hindi):

> You are being informed that electricity payment and other payment against your consumer ID: [IVRS No] amount of Rs. [X] are pending till month end [date]. This has been informed to you earlier. Moreover, there is also a notice issued against the above mentioned consumer ID. You are being informed that if the pending electricity payment of Rs. [X] is not paid within 15 days of the issue of this notice, then under electricity Act, 2003, Section 56 [...] the electricity will be disconnected once the notified period comes to an end. Please keep in mind

> that minimum electricity charges will be applied till the entire pending amount is not paid in full. In the case that the electricity amount is not paid, you will be charged under Section 138, as a punishable offence.

Though these notices are a part of standard utility practice, they are typically delivered in person rather than via post. Thus the main difference between this treatment arm and the status-quo is the mode of delivery, which eliminates any additional communication between the linesman and the consumer. In particular, this reduces the ability of consumers and linesmen to strike side-deals on receipt of a disconnection notice and reduces the costs of delivering notices by hand which may be otherwise substantial.

Most households receiving such notices are not actually disconnected after 15 days in the status quo. Treatment Arm E will only test the effect of the notice on payments and does not envisage follow up disconnections.

F. In-person disconnection notice, no disconnections: 2,500 consumers will receive a disconnection notice (identical text to the above), delivered in person by a utility employee. Payments made in response to this notice will be tracked over approximately the subsequent 30 day period (the exact duration might differ slightly depending on when the utility updates billing records). However, regardless of payment status, utility employees will *not* carry out a disconnection for households in this treatment arm.

G. In-person disconnection notice, subsequent disconnection: 2,500 households will receive a disconnection notice (identical text to above), delivered in person by a utility employee. Payments made in response to this notice will be tracked over approximately the subsequent 30 day period (the exact duration might differ slightly depending on when the utility updates billing records). Note that this duration is generous, exceeding the minimum payment time listed on the notice and allowable under existing regulation. After this period utility employees are expected to disconnect all consumers who have not paid off their arrears in full. Disconnected consumers are always able to regain their electricity access by paying off their full balance and a nominal connection fee of INR 200 (slightly less than 3 dollars).

This set of experimental arms enables us to measure the impacts of a series of enforcement actions ranging from very low-cost and light-touch (SMS messages) to high-cost, high-intensity enforcement (in-person notice with disconnections). Of this portfolio, the higher-cost and more punitive measures are the only options available in the status-quo, and are standard tools used by, and available to, distribution utilities in both developed and developing countries.

2 Data collection and outcome variables

We will collect outcome data from two sources: administrative data from the distribution company, a baseline survey conducted over the phone, and an endline survey after approximately 6 months. The endline survey is expected to also be a phone survey due to the COVID-19 pandemic but conditional on field conditions and budgetary constraints it may be converted to an in-person instrument. Access to utility administrative data is governed by a data use agreement between the discoms and the University of Chicago Trust in India.

2.1 Measuring treatment fidelity

We will observe data detailing treatment activities: when SMSes and letters are sent to each household, as well as records of each time a utility employee visits each household, and the date of disconnections and re-connections. This information allows us to infer when a household is *legally* without power. It is possible that disconnected households resort to theft - indeed this is a risk of disconnections - and we will seek to infer this through the phone survey by asking questions about consumption and appliance use. The treatments described above may also not be perfectly implemented by the utility in the field. These measures will thus allow us to precisely measure any deviations from treatment assignment. We discuss estimating ITT and LATE effects in the case of imperfect compliance in Section 3 below.

2.2 Administrative data

We expect to be able to collect administrative data spanning a baseline period, the period of the experimental implementation and extending to least four months of post-period data per experimental household. Households receive electricity bills on a regular billing cycle – typically 2 months in duration. For each billing cycle in the sample period, we will observe all features of the electricity bill, including the start and end date of the billing cycle and bill due date. The primary outcomes of interest are listed below

Payments: The billing data provides the monthly payment records for each household.

Electricity consumption: Metering records provide the household electricity consumption in KWh and the customer's total monthly bill in INR.

Outstanding arrears: The electricity bills also report outstanding arrears for each billing cycle. We can use this to calculate the ratio of unpaid amount to amount due in each cycle.

Collecting these data for at least six months after the experiment begins will allow us to estimate both short-run and medium-term treatment effects; we will extend our estimates beyond six months if given access to the data.

2.3 Baseline survey

We conducted a baseline survey by phone ending in December 1st, 2021. We contacted 19,000 customers from our 30,000 person experimental sample, and successfully completed 3,103 surveys. We included all customers in treatment group G (in-person disconnection notice, subsequent disconnection) in the survey sample, and then randomly drew 16,500 additional households from the remaining experimental groups. In this survey, we collected data on household demographics, electricity consumption and payments, and income. The survey instrument is provided in the Appendix.

2.4 Endline survey

We are planning to conduct an in-person endline survey to capture additional impacts of utility enforcement on household well-being. We will file an amendment to this PAP prior to conducting the endline survey that contains the details on the sample and our outcomes of interest.

3 Analysis

There are two sets of outcome variables about whom we are concerned when investigating the impact of the treatment interventions. The first set comes from utility administrative data viz. payments and consumption. In addition to treatment effects on these outcome variables we also wish to characterize the *heterogeneity* in treatment response. This is because an important goal for the utility is to learn more about how to target different enforcement actions, to maximize revenue recovery.

Beyond these administrative measures we are also interested in another set of outcomes to be measured using an endline survey. These outcomes are not otherwise observed or tracked by the utility. A full list of this second set of latter outcomes will be defined in an amendment to this PAP once the endline survey instrument design is completed. The outcomes we plan to measure in the endline survey include household expenditures, loans and savings, and a psychological well-being index.

3.1 Estimating treatment effects

We will estimate a set of analogous specifications as described by Equation 1 below for the main administrative data outcomes (electricity consumption, amount paid, and outstanding arrears) as well as for survey outcomes planned for endline (likelihood of enforcement; consumption expenditure; food expenditure; loan amount; and a psychological well-being index, made up of standardized versions of the three well-being questions).

For the administrative data outcomes, we will average outcomes over four months of post-period data (that is to say, two billing cycles following the relevant intervention date). In all regressions on administrative data outcomes, we will include average pre-treatment levels of the outcome variable, averaged over four months preceding the experiment start date.

When using survey data we will control for pre-treatment consumption and arrears. Owing to pandemic restrictions the baseline survey has a limited number of questions and may have a low response rate. We will therefore run specifications without baseline controls for the survey outcomes as our primary model and introduce baseline controls as additional specifications when possible.

$$Y_{is} = \beta_0 + \beta_1 T_i^B + \beta_2 T_i^C + \beta_3 T_i^D + \beta_4 T_i^E + \beta_5 T_i^F + \beta_6 T_i^G + \alpha_s + X_i + \varepsilon_{is} \tag{1}$$

where Y_{is} is an outcome variable for household i in stratum s; the treatments T^k, $k \in \{B, C, D, E, F, G\}$ are as described in Section 1.2, and indicators reflect *assignment* to treatment; α_s are strata fixed effects; X_i are controls for pre-treatment values, and ε_{is} is an error term. This main specification compares each of the SMS, mail, and in-person treatments to the control. In the main specification we will first allow for the BAU group to differ from the pure control by estimating β_1. We anticipate β_1 will not be statistically different from zero because status-quo enforcement activities are very limited and therefore a well implemented experiment may redirect enforcement resources such that the explicit control and the business as usual with no requests, are practically similar. If we find β_1 to be insignificant we will re-estimate this model without this term (pooling control and BAU groups) and use the resulting estimates of treatment effects as our main results.

3.2 Heterogenous treatment effects

Because we are particularly interested in optimal targeting, estimating heterogenous treatment effects is a central component of our analysis. We will use two approaches. In a first "traditional" approach, we will pre-define which variables can be used for targeting. This approach takes advantage of the fact that a utility has access only to a small amount of administrative data in the status quo, and recognizes that simple targeting rules are more feasible to implement.

To be precise we will use our experimental strata to trace out heterogeneity. Because we have many strata, we will first aggregate these up to a smaller number of groups defined by division, binned pre-treatment arrears, and binned pre-treatment consumption. A division is a geographic unit defined by the utility that is larger than a DC, but smaller than a circle. This yields a total of 42 groups.

To estimate within-group impacts, we will estimate Equation 1 (or its pooled variant that drops the BAU dummy) on the corresponding sub-samples.

In the second approach to quantifying heterogeneity, we will implement the approach described in Chernozhukov et al (2020).[4] This approach allows us to use machine learning methods to generate a proxy predictor for the treatment effect under each intervention and to sort units into different impact groups based on an average predicted treatment effect within group. Once this classification is complete it is possible to examine covariates associated with units with high or low predicted treatment effects and test to see if there are observable characteristics that are strongly and systematically correlated with predicted treatment effects. The main benefit of this approach is that we do not need to prejudge the dimensions along which we look for differential responses. The main cost is complexity and the fact that although this approach might describe heterogeneity effectively and with minimal assumptions, it need not lend itself to developing a targeting rule that is practically usable by policymakers.

3.3 Treatment dynamics

In addition to estimating the average impacts of our treatments, we are interested in testing whether our treatment effects vary over time. This is relevant to understanding the effectiveness of different treatments, as households may alter their behavior in the short run for some treatments and alter their behavior in the longer run for others. For this group of specifications, we only include administrative data, as we only have one survey wave.

To test for dynamic effects, we will estimate Equation 1 (and its variant that pools the BAU and control groups and drops the BAU dummy), interacting the treatment dummies corresponding to $\beta_2, \beta_3, \beta_4, \beta_5$, and β_6 with indicators for every 2 month period of post data. This produces treatment effects corresponding to these 2 two month windows. We expect to have at least 3 such coefficients, assuming a 6 month period of data collection in the post period. The regression is of the form described below:

$$\begin{aligned} Y_{is} = &\sum_{t=1}^{3} (\beta_2^t T_i^C + \beta_3^t T_i^D + \beta_4^t T_i^E + \beta_5^t T_i^F + \beta_6^t T_i^G) \times \mathbf{1}[\text{period} = t]_t \\ &+ \sum_{t=1}^{3} \mathbf{1}[\text{period} = t] \\ &+ \alpha_s + X_i + \varepsilon_{is} \end{aligned}$$

4. Victor Chernozhukov, Mert Demirer, Esther Duflo, and Iván Fernández-Val. 2020. *Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India.* NBER Working Paper No. 24678

where $t = 1$ is the two months after treatment, $t = 2$ is the next two months, and $t = 3$ is the final two months of the study. Y_{is} is an outcome variable for household i in stratum s; the treatments T^k, $k \in \{B, C, D, E, F, G\}$ are as described in Section 1.2; α_s are strata fixed effects; X_i are pre-treatment levels (for administrative outcomes), and ε_{is} is an error term. This main specification compares each of the SMS, mail, and in-person treatments to a pooled group containing both the pure control group and the BAU group; as above, if the BAU group and the pure control group are similar, we will pool them in additional specifications. To the extent that we are able to obtain pre-treatment data from the discoms, we will also include terms for earlier periods (i.e. allowing $t < 0$).

3.4 Local average treatment effects

The analysis described so far returns intent to treat (ITT) effects via Equation 1. These are relevant for the purpose of targeting rules because in a setting with imperfect implementation, the optimal utility approach should use the ITT and not the effect of the treatment on the treated (TOT). However it is of independent interest to estimate the TOT and we will do so by estimating a variant of Equation 1 using dummies for treatment status instead of treatment assignment. For example we would define an indicator for *actually* being disconnected versus being assigned to the disconnection treatment. Treatment status is endogenous since compliance may depend on household unobservables that affect the outcome. We will therefore use treatment assignment as an instrument and estimate this version of Equation 1 using a two-stage least squares approach.

3.5 Multiple hypothesis testing

We have only a small number of outcome variables from the administrative data: bill payments, remaining arrears, electricity consumption, and reconnections (if applicable). We therefore do not currently plan to employ any multiple testing adjustment. When we amend this PAP to include endline survey outcomes, we will discuss how we group these outcomes for a FWER/FDR-style adjustment.