Fixed Costs, Spillovers, and Adoption of Electric Connections

Barron, Manuel and Torero, Maximo

Department of Economics University of Santa Cruz, International Food Policy Research Institute - IFPRI

27 February 2015

Fixed Costs, Spillovers, and Adoption of Electric Connections

Manuel Barron and Maximo Torero

This Version: Feb 27, 2015

Abstract

We exploit exogenous variation in the price of connection fees to study the process of adoption of for formal connection to the electric grid in northern El Salvador. This variation, generated by randomly allocating discount vouchers among households, also generated exogenous variation in the share of neighbors receiving a discount ("voucher intensity"). We find that discount vouchers accelerated demand for formal connections, especially among households with an informal connection at baseline. We provide evidence that voucher intensity did not crowd out formal connections by facilitating informal access to the grid. Finally, we show that the electric utility could increase its customer base, revenue flows and profits by sharing part of the connection fee in the early years of an electrification project.

JEL codes: O33, O13

1 Introduction

This paper studies the determinants of adoption of electric grid, paying particular attention to the main suspect, adoption costs, as well as to one of the principal drivers of adoption in developing settings, spillover effects, which in our case arise as a combination of social learning and imitation. We then use these findings to illustrate a model in which it is profitable for the electric company to share part of the adoption costs with its potential customers.

There is little empirical evidence on the determinants of electricity adoption in developing countries.[1] Foster and Rosenzweig (2010) propose the lead determinants of technology adoption to be learning, profitability, incomplete markets, and "behavioral" mechanisms. In this paper we focus on the role of learning (in particular, social learning and imitation) and profitability (through variation in costs). Besides cost, we study spillovers from neighbors' adoption decisions, induced by the wealth of evidence generated by the literature on technology adoption in developing countries (e.g. Conley and Udry, 2010; Bandiera and Rasul, 2006; Kremer and Miguel, 2007; Dupas, 2010).[2]

To our knowledge, Bernard and Torero (2013), is the first empirical study on the adoption of electric connections. Using a similar identification strategy as this study, they find that the probability of connection increased among voucher recipients. In their study, households also responded to the number of number of vouchers allocated to the household's neighbors, something they attribute to "preferences interactions" (Manski, 2000), also known as "bandwagon effects". The main difference between this study and Bernard and Torero (2013) is that our study allows to trace out the dynamics of adoption by using multiple follow-up surveys. This proves to be key in our findings: while high-discount vouchers had a larger take-up than low-discount vouchers in the first follow-up, the effects even out starting on the second follow-up survey. We also find some evidence of spillovers, but the significance of the coefficients depends on the particular model. In addition, these coefficients are valid for the saturation experienced in our study, and are not generalizable if the whole population

[1] See Woolf (1987) for a study of adoption in early twentieth century USA.

[2] Conley and Udry (2010) study technology adoption among Ghanaian pineapple farmers and show evidence of a special type of social learning: Farmers will adopt based on the experiences of their unexpectedly successful peers. In turn, Bandiera and Rasul (2006) study how adoption of sunflowers varied by the farmer's network structure, finding an inverted U-shaped relationship. Studies on adoption of health technologies also show evidence of social learning. Kremer and Miguel (2007) in a follow-up to Miguel and Kremer (2004), find that students who had contacts exposed to deworming pills were less likely to take the pill, mainly because the students had learned about the externalities of getting dewormed. In a study on adoption of bednets, Dupas (2010) also finds evidence of social learning: households that did not receive the initial subsidy were more likely to purchase a bednet if they lived near other households that received favorable subsidies. The same study shows that willingness to pay for a bednet was higher among subsidy recipients than in the control group, evidence that households were also learning from their own experience.

received vouchers.

Vouchers increase adoption through several potential mechanisms: reducing connection costs, ameliorating credit constraints, providing incentives not to procrastinate, and increasing information about the program. In section 4.3 we argue that in this case the most likely mechanism is cost reduction.

Next, we draft a model showing that under some conditions the electric utility may profit from a voucher-allocation policy, de facto sharing connection costs with its customers. The rationale is that vouchers accelerate connection take-up, which increases revenues in the early years of the electrification project. As usual, equilibrium is achieved when the marginal gain of an increase in the subsidy offered equals the cost of subsidizing all clients, including those that would have connected without a subsidy. Using data from our experiment, we show that the utility firm would benefit from sharing part of the connection costs. Furthermore, the utility would increase its revenues by sharing 20% of the connection cost, but not by sharing 50%.

The rest of this paper is organized as follows. Section 2 drafts the theoretical background, while section 3 details the econometric approach. In section 4 we analyze the results, describe the population of adopters, and explore implications for infrastructure financing. Section 5 concludes.

2 Theoretical Background

2.1 The Household's Adoption Decision

We start from a static model with the basic assumption that households adopt an electric connection if the benefit b_i is larger than the cost c_i, subject to their budget constraint:

$$b_i > c_i \tag{1}$$

Where

$$b_i = f(A_i, \overline{E}_i, X_{0i}) \tag{2}$$

$$c_i = V_i + F_i \tag{3}$$

A_i is the vector of electronic appliances the household expects to acquire once they get on the grid.[3] This is key in that electricity provides utility gains only through the use of electronic appliances like light bulbs, television sets, refrigerators, and so on. Put it another

[3]Some households may already own, for instance, TV sets that are operated with car batteries.

way, these appliances are complement goods to electricity. $\overline{E}_i$ is the electrification rate around household i, to account of spillover effects as found in Bernard and Torero (2013). These spillover effects may arise from social learning, if households learn of the benefits of the appliances from their neighbors, or imitation, which could be a way of learning itself, or could be a reflection of preferences interactions ("bandwagon effects" or households keeping up with their neighbors, Manski, 2000). Finally, X_{0i} is a vector of demographic characteristics that may affect the benefits obtained from electrification (e.g. age composition, literacy, etc.).

On the other hand, the costs of electrifying a dwelling, once it is within reach of the grid, are a composition of variable and fixed costs. Variable costs cover wiring, installation of light bulbs and sockets, as well as potential upgrade in wall materials[4]. Fixed costs consist of, for example, any fees households have to pay as part of the application process, plus costs in time and effort in such process. We exploit the existence of such fees in our study setting to experimentally vary the cost of connection faced by each household in our sample, de facto creating exogenous variation in F_i. This exercise also creates exogenous variation in $\overline{F}_i$, the connection cost faced by $i's$ neighbors, which in turn generates exogenous variation in $\overline{E}_i$, the connection rate in household i's vicinity. We examine the role of both sources of variation to study adoption of grid connections.

Before we move forward, we consider other variables that may affect adoption of grid connections and their interaction with F_i and $\overline{E}_i$. Income is one of the most commonly cited determinants of technology adoption. Richer households are more likely to adopt an electric connection for two main reasons: first, the marginal disutility of paying F_i decreases with income, and second, since richer households can expect to buy more electronic appliances that are complement goods to grid connection, they can extract more utility from a connection. However, the way changes in F_i and $\overline{E}_i$ interact with income is not straightforward. Let's start with F_i. There is an income threshold m_1 above which F_i is trivial, so adoption rates will be unaffected by changes in F_i if income m_i is greater than this threshold. Similarly, there is an income threshold m_0 under which virtually any positive F_i is unaffordable. Hence only households with incomes $m_0 < m_i < m_1$ will respond to changes in F_i. Similarly, $\overline{E}_i$ will affect households within an income range, but it is not expected to affect neither very poor nor very rich households.

Second, we discuss the components of V_i. Wiring costs are an obvious component of V_i, and depends on the number of connections, sockets, light bulbs; but importantly also depends on the material on the walls. A dwelling cannot be electrified if its walls are made, for example, of wood, grass, or metal sheets; only if it has materials like cement, brick or adobe. Some households will need to upgrade their materials in order to pass the safety

[4] X% of households in our sample had "inadequate" walls, meaning...

inspection, which is not a trivial investment[5]. V_i includes time and effort for paperwork, hiring contractors and all activities related to getting the connection. If households exhibit hyperbolic discounting, they can postpone technology adoption indefinitely (Duflo, Kremer, and Robinson, 2011). To ameliorate this concern the validity of the discount had a time limit after the date of issuance, thus providing a salient benefit of taking action before its too late. V_i also includes the cost of solving any credit or liquidity constraints faced by the household. If solving liquidity or credit constraints in $t > 1$ periods is cheaper than doing so in one period, V_i falls with time, so the probability of adoption should increase as time passes, up to its equilibrium level.[6]

Before the program started, households that wanted a connection had to pay for the grid extension themselves, by paying the electric utility for posts, cables, etc. Although there is no official data on the costs, these are presumably high. Once a household got a grid connection, they could offer informal connections to their neighbors. Informal connections consist of a series of cables connected to a (formally connected) neighbor's outlet. This type of connections usually is enough to operate reliably a couple of light bulbs and at most a TV set. It can be argued that households with informal connections value electricity more than those that remained off-grid, since they have gone through the trouble of getting these connections in the first place. Additionally, most of these households already own some electronic appliances and thus have higher potential gains from adopting formal electricity than those with no access at all. In the empirical setting we account for this type of connection with an ordered choice model.

Two other factors that could affect adoption are characteristics of the household head and property rights. In Barron and Torero (2014) we show that males and females obtain different benefits from electrification, so we may expect them to behave differently. Females have higher valuation for health and education, and have more to gain from refrigerators and other kitchen appliances, while males have no such gains. Property rights may also affect the decision to connect to the grid, as they have been shown to affect other types of investments. If households lack property rights, they could be evicted and the investment would be lost. In fact, electrification may increase the probability of eviction since it increases housing value.

[5] On the other hand, a strong desire to get electricity may act as a nudge for households to update their construction materials. Now there is a salient benefit of upgrading. Before, there was no such benefit (despite the actual benefit of upgrading may be high).

[6] The pass of time could also lead to reductions in the number of connections, since some households may decide to drop (given e.g. unexpected prices), but this is not the case in our empirical setting.

2.2 Implications for Infrastructure Financing

In this section we study the conditions under which the utility would benefit by sharing connection costs with the beneficiaries, by increasing its client base early on in the project. Note that the electric company is a natural monopolist that cannot set price, which is determined by the regulator, or costs. However, we argue that the electric company could increase its customer base by paying a fraction of the connection fee. This practice is not uncommon, e.g. among mobile phone companies even in some developing countries, that pay a fraction of the phone in exchange for a commitment of use.

Let household benefits from electrification be given by b, which follows some density function $f(b)$. In this model, households will connect if $b > 0$, and will not connect otherwise.

Profits are given by

$$\Pi = \sum_{i=1}^{N} T_i(x)\left(R(q_i)q_i - x\right) - C_F \tag{4}$$

N is the total number of households within reach of the grid, $T_i = 1$ if household i is connected to the grid, 0 otherwise, $R(q_i) = p(q_i) - c(q_i)$ is the gross profit obtained from household i (gross of fixed and connection costs), q_i is the quantity of electricity consumed, x is the part of the connection fee paid by the firm, and C_F is the fixed cost of grid extension.

At each round households can be classified as follows: Always-Takers, Never-Takers, and Compliers. *Always-takers* are households that would have connected even without the discount. Next, *Never-takers* are households that would not connect even if they received the discount. Finally, *Compliers* are households that would connect if and only if they receive the discount.

Note that the status of a household as a complier, never-taker or always-taker, is contingent on the period. In particular, a household may be a complier at round t and an always-taker at round $s > t$. For instance, take a household that without the voucher would have decided to connect in period 3, but with the voucher decides to connect in period 2. This household is a complier in period 2, but an always-taker in period 3. Despite its not possible to know which household is a complier, the size of the complier subpopulation is given by the estimates of the β coefficients in the adoption equations (section 3.1).

Formally, x has and effect on the extensive margin (through new customers), as well as on the intensive margin (more consumption by the existing customers). To simplify matters, we assume the intensive margin to be zero. This simplifies the algebra at little cost in terms of insights. This assumption is perhaps more appealing if we think of intertemporal extensions to this model, where lifetime electricity consumption would vary little by receiving a $20 or $50 discount on the connection fee.

We set the equilibrium such that marginal cost equals the marginal revenue. Assuming no effect on the intensive margin (i.e. that q_i does not depend on x), this simplifies to:

$$\sum_{i=1}^{N} \frac{\Delta T_i}{\Delta \alpha}(R(q_i)q_i - x) - \sum_{i=1}^{N} T_i = 0 \tag{5}$$

Note that $\frac{\Delta T_i}{\Delta \alpha}$ can take only one of two values. It is zero for always takers and never takers, and one for compliers.In this context, always-takers are households with $b > 0$, compliers are households with $-x > b > 0$ and never-takers are households with $b < -x$.

$$\frac{\Delta T_i}{\Delta \alpha} = \begin{cases} 1 & \text{if } -x < b < 0 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

To keep simplifying matters, assume that q_i is constant among compliers. Thus,

$$R(q_i)q_i = Rq \ \forall \ \text{i} \in \text{C} \tag{7}$$

With this, we can re-write (6) as follows[7]:

$$Rq \int_{x}^{0} f(b)db = x \int_{x}^{0} f(b)db + \int_{0}^{\infty} f(b)db \tag{8}$$

Equation (8) means that the additional gains the utility obtains from each complier at this new level of x must be enough to pay x to each of those compliers, plus the marginal subsidy increase to each always-taker (those who would have decided to connect before the incremental change in the subsidy). This confirms the initial insight that if compliers generate enough gains to the utility to pay for the additional subsidy to the always-takers, it is profitable for the utility to offer a strictly positive subsidy.

We show this in the case of normally distributed b. If $b \sim N(\mu, \sigma)$, from (8) it can be shown that:

$$\left[\Phi\left(\frac{x^* + \mu}{\sigma}\right) - \Phi\left(\frac{\mu}{\sigma}\right)\right](Rq - x^*) = \Phi\left(\frac{\mu}{\sigma}\right) \tag{9}$$

, which implies that if $Rq > x^*$, then $x^* > 0$.

In order to find an upper bound to x^*, note that the model discussed in section 3 implies an inverse relationship between q_i and x (Zilberman and Liu, 2011). Marginal consumers,

[7] We have omitted an intermediate step. To wit,

$$Rq \times N \Pr\left(\frac{\Delta T_i}{\Delta x} = 1\right) = x \times N \Pr\left(\frac{\Delta T_i}{\Delta x} = 1\right) + N \Pr(T_i = 1)$$

those who need a discount to connect, will consume less than the average consumer. The larger the required discount, the smaller q_i. Thus, there is a threshold value for x, call it x_T, such that $\forall x > x_T$

$$Rq \int_{x}^{0} f(b)db < x \int_{x}^{0} f(b)db + \int_{0}^{\infty} f(b)db \tag{10}$$

This simple model can be the base of a dynamic optimization problem in which the firm either subsidizes in the first period only or subsidizes at every period. In its current form it shows some interesting insights. If the compliers can pay for themselves ($Rq > x$), then the optimal discount is positive. To pin down the optimum discount, we would need to impose some more structure on the problem. Since this paper is not about the optimal subsidy, we leave this issue for future research.

There are many reasons why this type of arrangements may not be implemented widely. The most obvious reason would be due to lack of knowledge of $f(b)$. Next, it could also be the case that the electric company may face the probability of default in payments. Note that the simple model sketched in this can incorporate the probability of default by multiplying the left hand side of (8) times the probability of payment. However, more sophisticated modeling of the probability of default and its associated costs could provide more interesting insights. Third, it is the case that someone has to come up with this solution, and it is possible that the electric companies have not done so yet.

3 Empirical Approach

3.1 Standard Approach

Our main estimating equation is given by:

$$conn_{it} = \beta_0 + \beta_1 F_i + \beta_2 \overline{E}_{it} + \beta_3 X_{i0} + \lambda_t + \varepsilon_{it} \tag{11}$$

To include household fixed effects and still be able to estimate β_1 we follow alternative methodologies. (i) we create a *post* dummy that takes the value of 1 in periods 2 through 5, and interact and interact all the explanatory variables with *post*. (ii) we interact each explanatory variable with the round fixed effects, thus allowing for different effects in different periods without imposing a linear or quadratic trend. (iii) we interact the explanatory variables with t and t^2. To allow for non-linear effects of F_i on connection, we also replace F_i by dummies corresponding to the \$20 and \$50 discounts.

Since $\overline{E}_{it}$ is endogenous, we employ two strategies to find its causal effect on $conn_{it}$. First, we use the average fee among household i's neighbors as an instrument for $\overline{E}_{it}$. Second, we replace $\overline{E}_{it}$ by the share of eligible neighbors around household i that received a discount voucher, $\bar{s}_i$. This estimation gives the "reduced form" coefficient.

We start with a deliberately simple approach to study electricity adoption, we start by estimating:

$$conn_{it} = \beta \text{ Voucher}_i + \gamma^{100} V_i^{0-100} + \lambda^{100} N_i^{0-100} + \omega' X_{i0} + \sum_{t=2}^{4} \nu_t + \varepsilon_i \tag{12}$$

, where $conn_i$ indicates whether household i has a formal connection in year t, Voucher indicates whether the individual received a voucher, V_i^{0-100} indicates the number of households that received a voucher within a 100 meter radius of household i, while N_i^{0-100}, the total number of eligible households around 100 meter radius of household i. X_{i0} are individual baseline characteristics. The ν_t are year fixed effects are included because the follow-up period includes multiple years; we collected up to three yearly measurements per observation in years 2 through 4 of the study. ε_i is the unobserved residual. As is usual in this approach, the standard errors will be clustered at the level of treatment, i.e., the household level.

Note that including the share of connections as an explanatory variable for i's decision to connect would be inadequate, since the neighbors connection decision depend upon i's decision. Hence, we only present this "reduced form" type of results.

Next, we allow for different effects by size of the discount, by estimating:

$$conn_{it} = \beta_{20} V20_i + \beta_{50} V50_i + \gamma_{100} V_i^{0-100} + \lambda_{100} N_i^{0-100} + \omega' X_{i0} + \sum_{t=2}^{4} \nu_t + \varepsilon_{it} \tag{13}$$

Third, we estimate the effect of the amount of the fee, as well as the average fee in the neighborhood of household i. This imposes some structure in the regression, namely, that costs affect connection rates linearly.

$$conn_{it} = \beta \times \text{fee}_i + \gamma_{100} \times \text{fee}_i^{0-100} + \omega' X_{i0} + \sum_{t=2}^{4} \nu_t + \varepsilon_{it} \tag{14}$$

The results from the above regressions will give the effects of discount vouchers on adoption averaged over the three follow-up surveys. To analyze the dynamic effects, we interact the right hand side variables with time dummies.

3.2 Informal Connections

A second issue we need to take into account is the existence of informal connections. To take into account the "essential heterogeneity" (Heckman, Urzua, and Vytlacil, 2006) at baseline between households with no connection and those with an informal connection, the above equations could be run separately for households with an informal connection and for those with no connection at baseline, or we could fully interact each variable in the right-hand-side variable with a dummy for informal connection at baseline.

We study switching patterns between three connection types: no connection, informal connection, formal connection. Reductions in F unequivocally increase the probability of household i getting formal connection to the grid, but reductions in $\overline{E}$ generate two opposing forces: on the one hand, imitation-type spillovers would increase the probability of getting a formal connection, while on the other hand, the more neighbors of i are formally connected to the grid, the easier it is for i to get an informal connection.

We study the probability of switching among different alternatives. For instance, we look at the probability of a household switching from no electricity to formal electricity at some point over the study period, ignoring households that had informal electricity at baseline. We conduct similar analysis for the probability of switching from no electricity to informal electricity, and from informal to formal. While this has the disadvantage of leaving aside part of the sample, it provides robust insights without imposting structure to the model.

Alternatively, we exploit the natural ordering among these three types and estimate and ordered probit model. In this setting formal connections (y=2) are the best type of connection, followed by informal connections (y=1), and finally by no connection (y=0). Ordered choice models start by assuming that households choose the connection type that maximizes their utility level. Subjects will choose their connection type y depending on the value of a latent, unobservable variable y^*, such that:

$$y^* = \beta x_i + \varepsilon_i \tag{15}$$

, where x_i is a vector of explanatory variables. Although y^* is unobservable, we observe y, the subject's choice:

$$y = 0 \quad \text{if } y^* \leq \alpha_1 \tag{16}$$

$$y = 1 \quad \text{if } \alpha_1 < y^* \leq \alpha_2 \tag{17}$$

$$y = 2 \quad \text{if } y^* > \alpha_2 \tag{18}$$

, for some (estimable) cutoff parameters α_i. If we assume that ε_i follows a normal distribution, we get to the ordered probit model. In this model, the probability of choosing each alternative is given by:

$$\Pr(y=0) = \Phi(\alpha_1 - x_i\beta) \tag{19}$$
$$\Pr(y=1) = \Phi(\alpha_2 - x_i\beta) - \Phi(\alpha_1 - x_i\beta) \tag{20}$$
$$\Pr(y=2) = \Phi(x_i\beta - \alpha_2) \tag{21}$$

and the marginal effects are given by

$$\frac{\partial \Pr(y=0|x_i)}{\partial x_k} = -\beta_k\phi(\alpha_1 - x_i\beta) \tag{22}$$
$$\frac{\partial \Pr(y=1|x_i)}{\partial x_k} = \beta_k[\phi(\alpha_1 - x_i\beta) - \phi(\alpha_2 - x_i\beta)] \tag{23}$$
$$\frac{\partial \Pr(y=2|x_i)}{\partial x_k} = \beta_k\phi(\alpha_2 - x_i\beta) \tag{24}$$

, where $P(y=j|x_i)$, $j=0,1,2$ denote the probability of choosing no connection, informal connection, or formal connection, respectively. Further, note that the direction of the effect of x_k on $P(y=0|x_i)$ and $P(y=2|x_i)$ is given by the sign of β_k, but this is not true for $P(y=1|x_i)$, since the effect depends on the sign of $\phi(\alpha_1 - x_i\beta) - \phi(\alpha_2 - x_i\beta)$. Note that the marginal effects for each observation depend on their particular x_i, so they will vary between observations. Thus the effect may be positive for some households and negative for others. This flexibility is especially interesting for us, since it is theoretically ambiguous whether spillovers will increase or decrease informal connections.

4 Results

Table 2 shows how changes in F_i and $\overline{E}$, empirically measured by "voucher" and "s100" affected connection rates. Column 1 pools the four follow-up surveys (survey rounds 2 through 5), showing that voucher recipients were on average 12 percentage points more likely to get a formal connection in the post-baseline period. Second, there appear to be important spillovers: a 10 percent increase in s100 increased the probability of connection by 1.3 percentage points. This indicates potentially large externalities, which analyze in greater detail later. The results in column 2, which adds household fixed effects to the specification in column 1, are identical, which clears any potential doubts on whether randomization worked in this relatively small sample.

In specifications equivalent to columns 1 and 2 adding s200, the coefficients are unchanged and s200 turns out not significant (online appendix). For the sake of parsimony we don't include s200 in our main specifications, but we include such specifications in the online appendix as robustness checks. Columns 3 and 4 replicate the first two columns interacting voucher and s100 with round dummies, which allows exploring time patterns. Although the differences are not statistically significant, the coefficients on voucher are larger in rounds 2 and 3, around 15 percentage points, than at rounds 4 and 5, around 10 percentage points. This decline owes to the fact that the connection rate in the non-encouraged group started catching up the encouraged group, as shown in Figure 1. The effect of s100 doesn't appear to have such a trend, which would indicate persistence in the externalities.

Table 2 analyzes take-up using a different explanatory variable: the inspection fee. The variable takes the value of 1 if households had to pay \$100, 0.80 if they received the 20% discount and 0.50 if they received the 50% discount. In this specification, spillovers are captured by the average fee paid by eligible neighbors in a 100m radius. Pooling together the three follow-up surveys shows that a \$10 reduction in the fee would increase the probability of connection by 2 percentage points. Once again, including household fixed effects leaves the point estimates unchanged.

Next, we go on to the dynamics. The coefficient on the inspection fee at round 2 is -.32, significant at the 1% level. By round 3 the effect size decreases to -0.26, significant at the 5%. The effect loses significance and approaches 0 in rounds 4 and 5 (-0.13 and -0.11, respectively). This indicates that a lower fee increased adoption early on, but not later in the process.

The average fee has significant effects on connection take-up, with a \$10 reduction increasing the probability of connection by 3.5 percentage points over the study period. Analyzing the dynamics we see that the effects are significant in all rounds, fluctuating from .29 (round 2) and .45 (round 3) te. This suggests that peer effects did not decrease with time. In column 5 we explore linear and quadratic time trends in the effects of fee and fee100. While this specification imposes more structure than the previous, said structure is consistent with the more flexible specifications in the previous columns.

Now we turn to analyze how $\overline{E}$ affected connections. As discussed in the main text, we use s100 (or fee100) as an instrument for $\overline{E}$. An additional connection within 100 meters increases the probability of i's connection by 10 percentage points (10.3-11.1, depending on the exact specification), almost the same effect as household i itself receiving a voucher. When expressed in percentage of eligible neighbors, we see that a 10 percentage point increase in the share of eligible neighbors that connect to the grid increased the probability of connection by roughly 2 percentage points (1.9-2.2, depending o the specification).

The exclusion restriction in this case requires that the instrument affects connection decision only through its effect on $\overline{E}$. Lets think about it. It means that s100 altered $\overline{E}$, which in turn affected E. The exclusion restriction would be violated if, for instance, i's neighbors who received vouchers would encourage household i to connect by providing them information about the connection process, making them aware of the electrification process, etc. If household i did not receive a voucher and their neighbor did, we would expect household i to refrain from connecting, maybe hoping to get a voucher at a later stage or being upset for not having received said voucher. The information story is not strong either, since grid electrification projects are easily visible by all in the community. If it were an information story we would have the externalities declining with time. We see no such time trend. On the contrary, we spillover effects are persistent across our study period.

4.1 Connection Choice

We now turn to the analysis of connection choice. As discussed previously, households in our sample have the option of getting an informal connection, which in the analysis so far has been counted as begin off the grid. We conduct two types of empirical approaches. First, we look at the probability of switching between any pair of connection types (from no electricity to formal, from informal to formal, from no electricity to informal[8]).

The dependent variable takes the value of 1 if the household switched at some point during the study period, and 0 otherwise. The sample only includes households that had no connection at baseline (i.e., it drops households with informal connection at baseline). This reduces the sample size to 275 observations. Once we control for covariates, the effect of voucher and s100 are not significant: households that did not have an informal connection did not respond to vouchers. The coefficient on s100 is too large to reject positive externalities, but it is not statistically significant. Columns (3) and (4) report the results of switching from informal to formal connections. Both voucher and s100 result strongly significant in this subsample. Among this group, voucher recipients are 15 percent more likely to switch to a formal connection, and having 40% of eligible neighbors receiving a voucher (mean s100) increased the probability of connection by 4 percentage points. Finally, columns (5) and (6) show that vouchers did not affect the probability of switching from no connection to an informal connection.

Next, we present the result of ordered probit estimation for connection choice. This method imposes more structure in the analysis in two main ways. First, it exploits the

[8]There are just 16 cases of households switching from informal to off-grid and no cases of households dropping their formal connections

difference in quality, which close to reality since differences in reliability and other characteristics make formal connections better than informal connections. Second, it requires to impose the assumption that the disturbances term follows a normal distribution. Table 6 shows the main results of the model.

Figure 3 plots the modeled probabilities of each type of connection by round, according to the ordered probit model. In panel (a) we estimate the probability of formal connection by round and plotted the kernel densities, to give a better sense of the evolution of connection rates. Each round the mean increases and the variance decreases. Panels (b) and (c) show that the probabilities of having an informal connection or being off-grid, respectively, moved in the opposite direction. Hence, informal connections don't seem to be crowing out formal connections. On the contrary, formal connections seem to be crowding out informal ones.

Figure 4 plots the marginal effects of discount vouchers on the probabilities of formal (panel (a)) and informal (panel (b)) connections, from the ordered probit model in Table 5 column 2. In this plot the marginal effect for each household was estimated, given their own characteristics. Panel (a) shows that the effect of receiving a voucher is concentrated around 0.10 at around 2, and that it fades away in later rounds, as non-recipients catch up in formal connection rates. The effect of receiving a voucher on the probability of having an informal connection is initially disperse, with a mean of around 0 and a range of -0.02 to +0.20, meaning that it increased the probability of an informal connection among some households by 2 percentage points and decreased it by 0.2 percentage points among others. In rounds 3 and later, the effect is negative among almost all households, and concentrated around -0.02. Plot (c) shows that voucher recipients were less likely to remain off-grid by round 2, but then the effect got diffused, as non-recipients adopted some type of connection.

Now, turning to spillover effects, Figure 5 plots the relationship between s100 and connection type. To obtain these figures, we estimated the model in Table 5, column 2, and obtained predicted values by simulating different values of s100 while keeping each household's observed characteristics. Both relationships are roughly linear, but the slope on the probability of having a formal connection is positive, while the slope on the probability of having an informa connection is negative. Panel (b) shows that the marginal effect of s100 on the probability of having a formal connection is positive but declines as s100 increases: household i's first connected neighbors have slightly higher influence on i's decision to get a formal grid connection, and while having more connections increases the probability of formal connection, it does so at a decreasing rate. On the other hand, the marginal effect of s100 on the probability of having an informal connection is small, negative, and fairly constant across values of s100.

4.2 Adopters

Next, we examine the characteristics on adopters to get better insights on who may be the beneficiaries of a cost-sharing policy. These are just descriptive regressions and we claim no causality, they provide insights on the characteristics of adopters. The dependent variable takes the value of 1 if the voucher recipient connected at some point over the course of the study and zero otherwise. This regression controls for subdistrict fixed effects.

Households with informal electricity at baseline were 18 percentage points more likely to take up the voucher, as were households with property title (8 percentage points) and households with floor other than dirt (10 percentage points). Households with more members were also more likely to adopt. Material on the walls had no apparent relation with adoption. Income is positively related to adoption, but the magnitude is economically small: a $1,000 increase in annual income is associated with a 1.4 increase in the probability of connection. Aged, gender and literacy status of the household head had no apparent relation with voucher take-up.

Clearly all the variables are related, but empirically, the main marker to identify households that will respond to vouchers is informal connections. Those households have gone through the trouble of getting the connection from their neighbors, hang cables, etc., which reveals they have higher valuation for electricity than the rest.

4.3 Mechanisms

Earlier we argued that vouchers increase adoption by reducing connection costs, ameliorating credit constraints, providing incentives not to procrastinate, and increase information about the program. In this section we analyze the empirical evidence to try to identify which of these channels likely played a larger role.

It doesn't seem like vouchers worked as commitment devices. If households needed incentives not to procrastinate, both type of vouchers should have similar effect at each round. On average they do, but the timing is not consistent with this story: high discount vouchers have a higher adoption rate than low-discount vouchers. It is easier for a household to come up with $50 than $80 for the connection fee, so households in the latter group will need a longer period to connect. In addition, the non-encouraged group had high connection rates starting from the second round (50%), up to 80% by the third round. This is consistent with vouchers reducing connection costs.

Although the survey did not include questions on information about the electrification program, it seems unlikely that vouchers increased program awareness. Projects of this type are easy to detect by households, since they require erecting posts and other types of construction work.The reimbursement provided by the discount vouchers could have been used to pay the cost of a loan, thus lifting credit constraints. However, access to credit (either formal or informal) did not increase among voucher recipients. This is not to say that these households are not credit constrained, but that vouchers did not increase access to credit markets.

4.4 Implications for Infrastructure Financing

Table 5 shows some calculations to simulate revenues per household under different subsidy schemes in the initial years of the electrification program. First, we have the electrification rate per year under each type of voucher, from Table 1. To get at the average revenues perceived by the electric utility we use the average bills for households with a formal grid connection. The average electric bill in round 2 was \$8.32 per month, so the average household generated \$99.84 of revenues per year.

Multiplying the connection rate times the average bill we find that the utility received \$49.92 per household in the area. If the firm shares 20% or 50% of the cost, the figures are \$60.90 and \$66.92, respectively. To find the most profitable alternative we subtract the connection costs that the company shared in each scenario. Note that the company pays \$20 for all connections in round 2 and for the increase in connection rates in the future rounds.

With this, we can estimate revenues under each of the subsidy schemes. Although this is not vital for the results, we assume a 5% interest rate to calculate the revenues in the first three years of operation. If the company provides no subsidy, it would receive roughly \$180 per household in the first three years. Sharing 20% of the connection costs would bump those revenues by 10%, up to \$200 per household. On the other hand, the 50% of the cost proves not profitable, since revenues would be \$175, below the no-subsidy scenario.

If the electric utility were able to discriminate, a more profitable, but more controversial, solution would likely be implemented. Remember that cost sharing is “repaid” via the additional revenues extracted from the monthly bills of compliers. If the company could identify a “high bill” group, it could target subsidies to these households and improve revenues further. One such group is better off households. As we saw previously, households above the median income responded to the vouchers, while poorer households did not. Also, the average electric bill in the former group is double that of the latter: better-off households spent on average \$11.02 per month, while poorer households paid \$5.22 for electricity. Since richer households are morel likely to respond to the voucher and to “pay back” the subsidy with higher consumption, profit maximization on the part of the utility requires offering subsidies only to the high-consumption group. Table 6 shows that the profit-maximizing strategy requires the utility to offer 20% discounts to better-off households and no subsidy to poor households.

5 Conclusions

This paper studies the role that connection costs and spillover effects play in the adoption of formal electric connections. We observe voucher recipients connecting 1-2 years earlier than the control group, but the control group catches up to them by the end of the study. Thus, in our setting households respond to time-limited reimbursements by adopting earlier than the counterfactual.

Second, spillover effects seem to play an important role in adoption, and their effects do not reduce down with time. An additional connection within 100m of a household increases the probability of that household connecting formally to the grid by 10 percentage points, roughly the same increase generated by vouchers.

Third, we analyze the possibility that households adopt informal connections, by modeling the probability of switching between different alternatives and by estimating an ordered choice model. The first strategy suggests that households with informal connections at baseline were much more responsive to vouchers and spillover effects, while these variables did not affect the adoption rates of households with no connection at baseline. The ordered choice analysis shows that vouchers (allocated directly or to neighbors) increased the probability of formal connections, and decreased the probability of adopting informal ones. This solves any concern about informal connections crowding out formal ones.

Among our treatment group, households with informal electricity at baseline were more likely to take-up the voucher, as well as households with property titles. Better-off households (no drifter, higher income) were also more likely to take-up the voucher. Despite gender differences in the benefits of electricity (see e.g. (Barron and Torero, 2014)), gender of the household head was not correlated to voucher take-up. Similarly, other characteristics of the household head like age or literacy status were uncorrelated with take-up.

Finally, we show that the electric utility can actually increase its revenues by providing discount vouchers in a similar fashion as in this study, given that vouchers increase the utility's customer base and revenue flows. In our study, the "optimal" strategy consists in paying $20 of the inspection fee. Since connection rates among low- and high-discount voucher recipients are very similar.

References

Bandiera, Oriana and Imran Rasul. 2006. "Social networks and technology adoption in northern mozambique*." *The Economic Journal* 116 (514):869–902.

Barron, Manuel and Maximo Torero. 2014. "Household Electrification and Indoor Air Pollution." *Mimeo* .

Bernard, Tanguy and Maximo Torero. 2013. "Bandwagon Effects in Poor Communities Experimental Evidence from a Rural Electrification Program in Ethiopia." .

Conley, Timothy G and Christopher R Udry. 2010. "Learning about a new technology: Pineapple in Ghana." *The American Economic Review* :35–69.

Duflo, E, M Kremer, and J Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and." *The American Economic Review* 101 (6):2350–90.

Dupas, Pascaline. 2010. "Short-run subsidies and long-run adoption of new health products: evidence from a field experiment." Tech. rep., National Bureau of Economic Research.

Foster, Andrew D and Mark R Rosenzweig. 2010. "Microeconomics of technology adoption." *Annu. Rev. Econ.* 2 (1):395–424.

Heckman, James J, Sergio Urzua, and Edward Vytlacil. 2006. "Understanding instrumental variables in models with essential heterogeneity." *The Review of Economics and Statistics* 88 (3):389–432.

Kremer, Michael and Edward Miguel. 2007. "The illusion of sustainability." *The Quarterly Journal of Economics* 122 (3):1007–1065.

Manski, Charles F. 2000. "Economic Analysis of Social Interactions." *The Journal of Economic Perspectives* 14 (3):pp. 115–136.

Miguel, Edward and Michael Kremer. 2004. "Worms: identifying impacts on education and health in the presence of treatment externalities." *Econometrica* 72 (1):159–217.

Woolf, Arthur G. 1987. "The Residential Adoption of Electricity in Early Twentieth-Century America." *The Energy Journal* 8 (2):19–30.

Zilberman, D. and J. Hsing-I Liu. 2011. *On Consumption Indivisibilities, the Demand for Durables, and Income Distribution*, chap. 52. Dynamics, Games, and Science I, 785–809.

Figure 1. Connection Rates by Treatment Arm

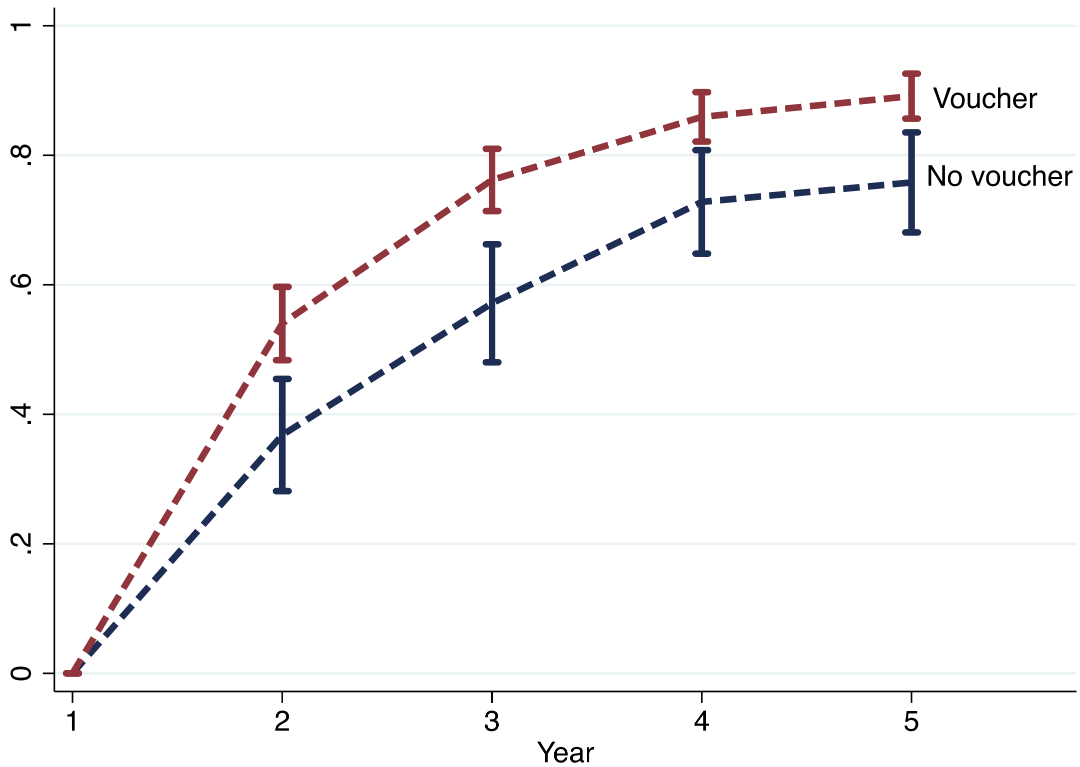


Figure 2. Connection Rates by Treatment Arm

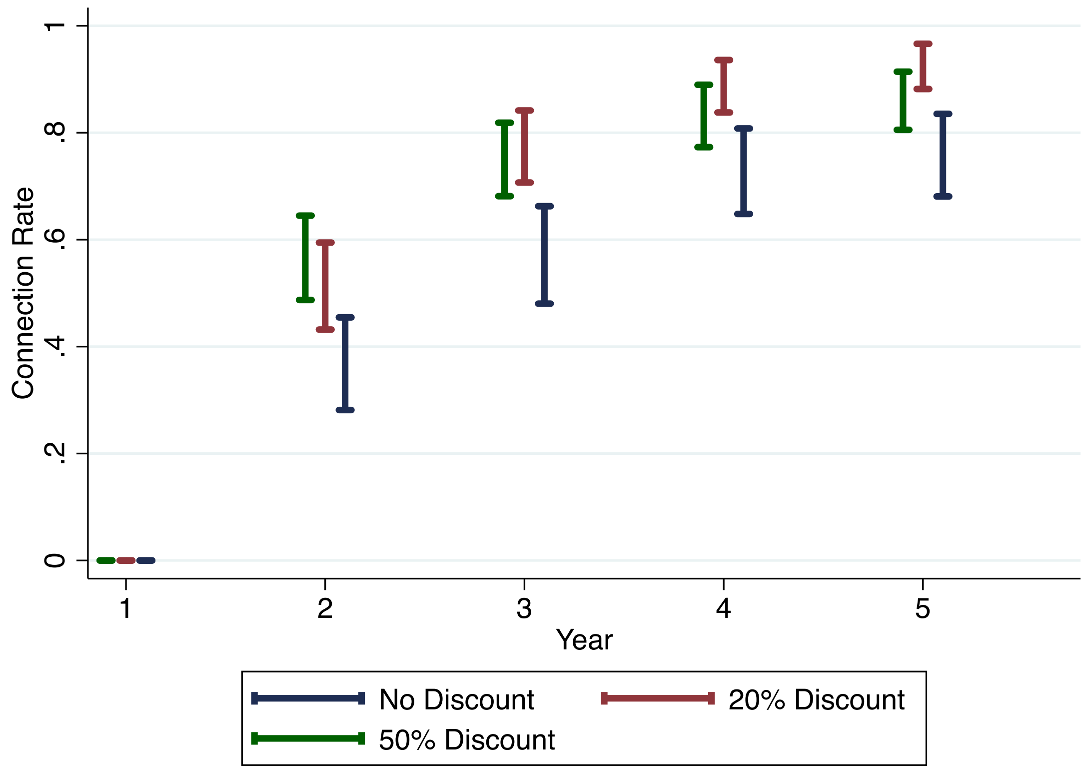
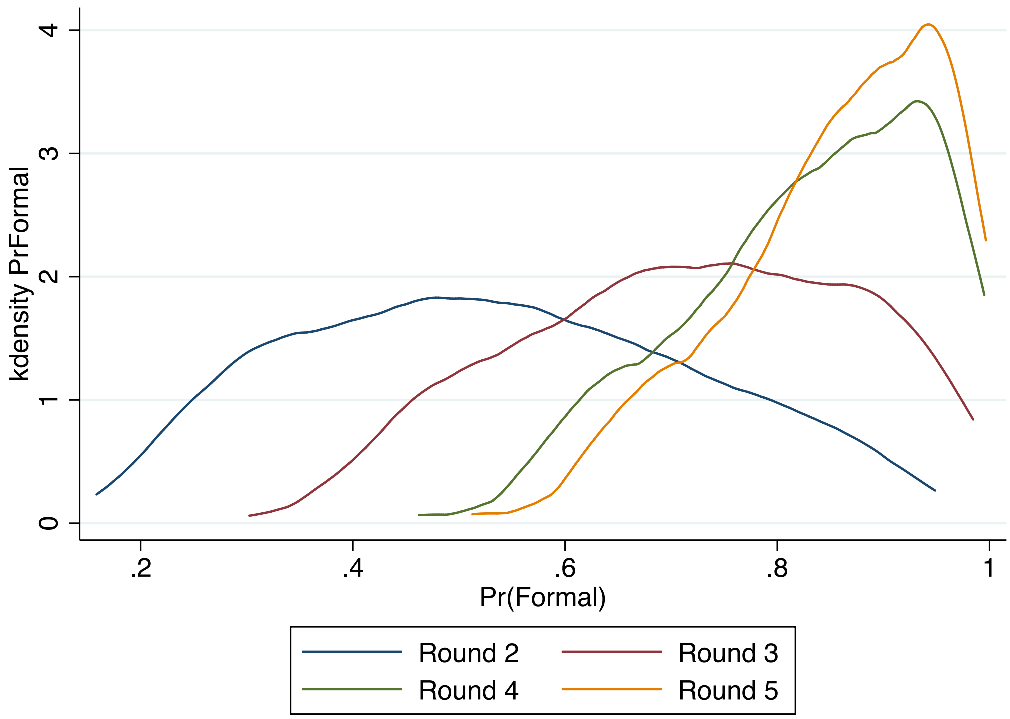
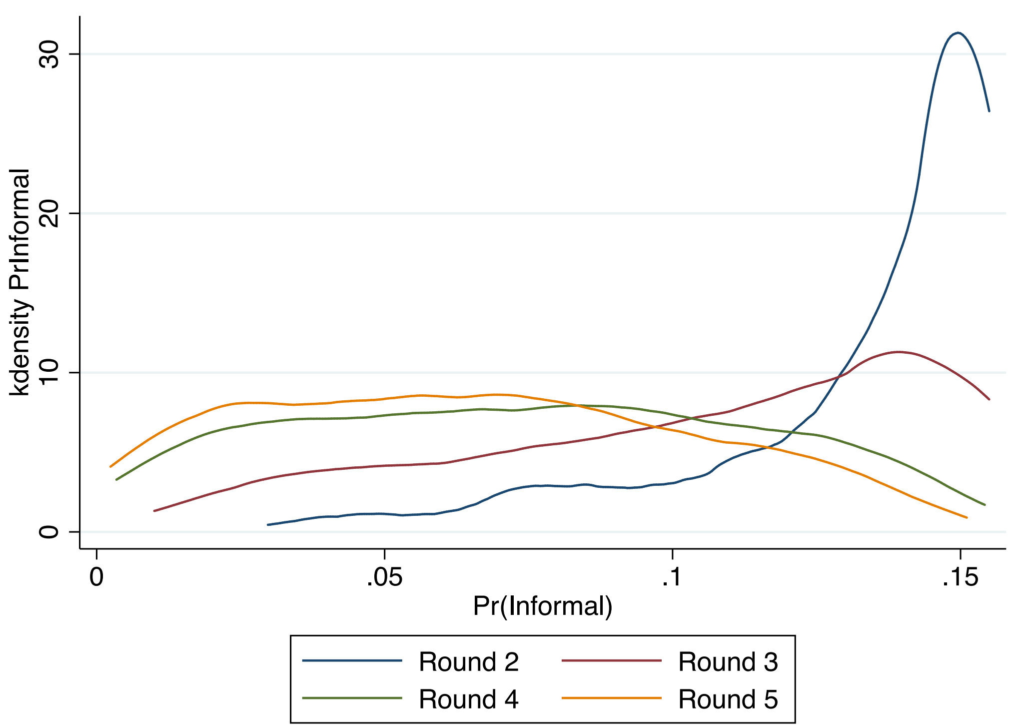


Figure 3. Connection Type by Round (Ordered Probit)

(a) Probability of Formal Connection



(b) Probability of Informal Connection



(c) Probability of No Connection

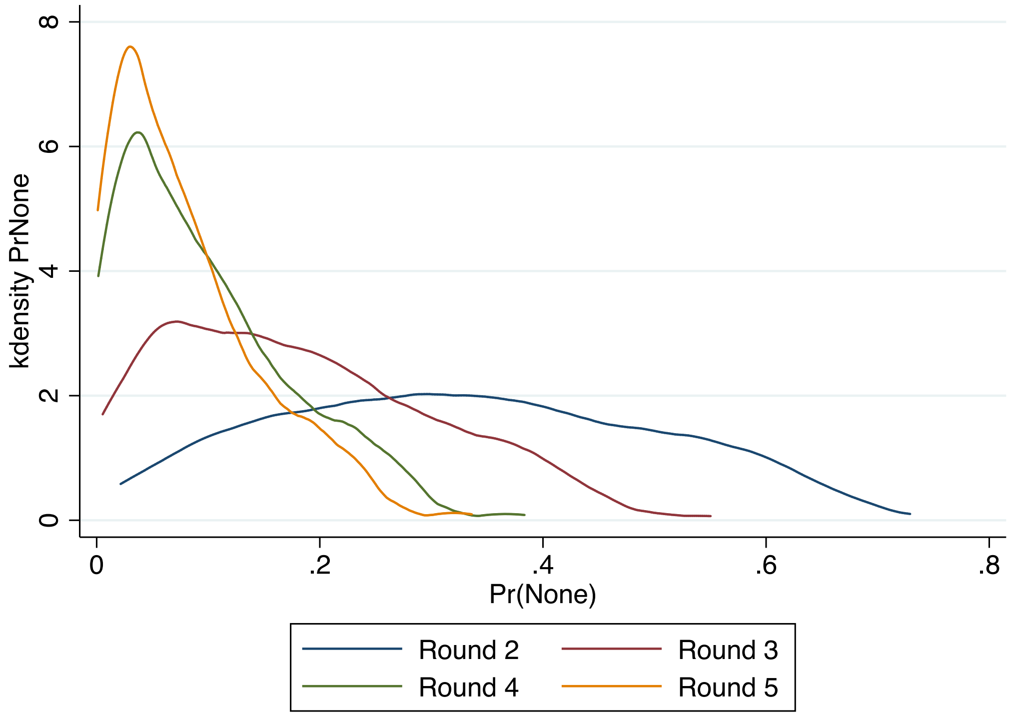
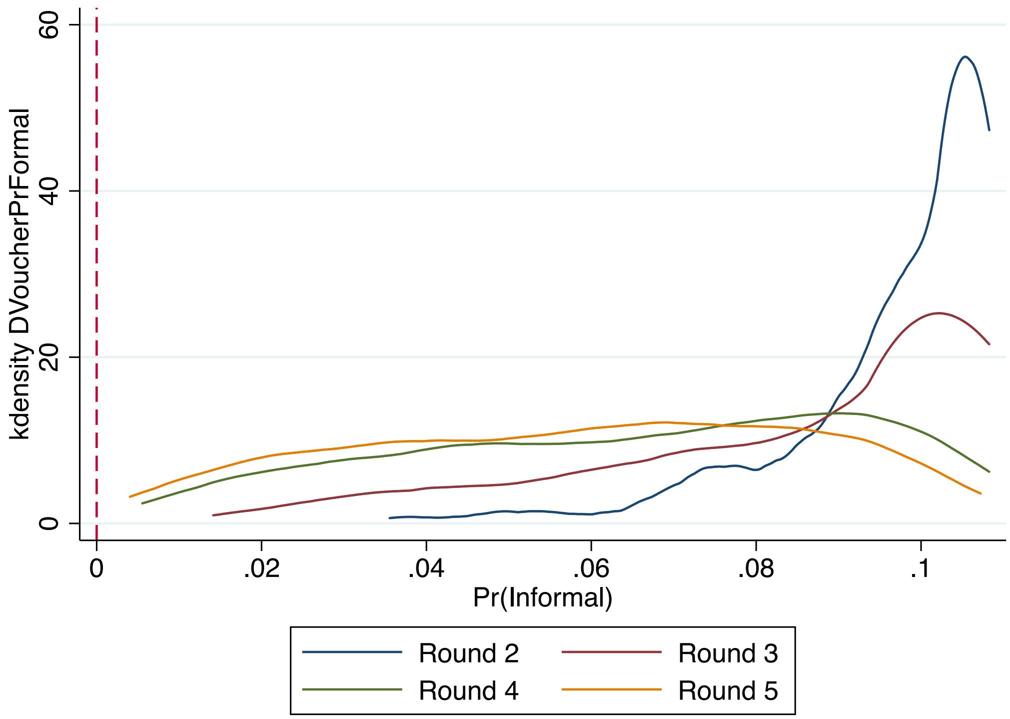
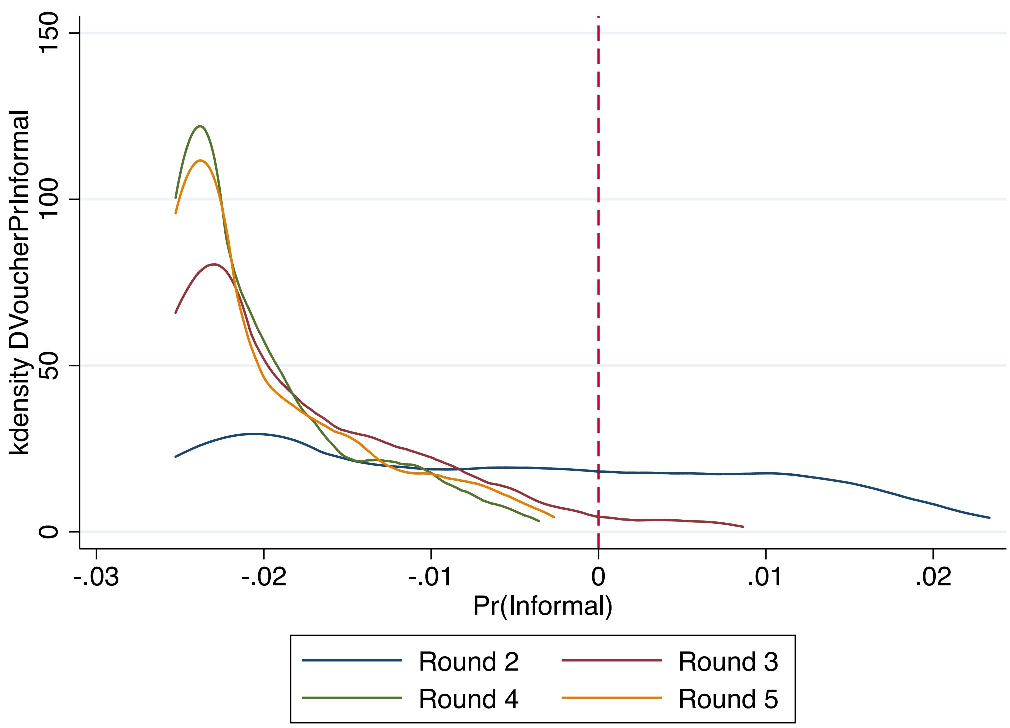


Figure 4. Effects of Voucher Allocation on Connection Type By Round (Ordered Probit)

(a) Effect of Voucher on Pr(Formal)



(b) Effect of Voucher on Pr(Informal)



(c) Effect of Voucher on Pr(No Connection)

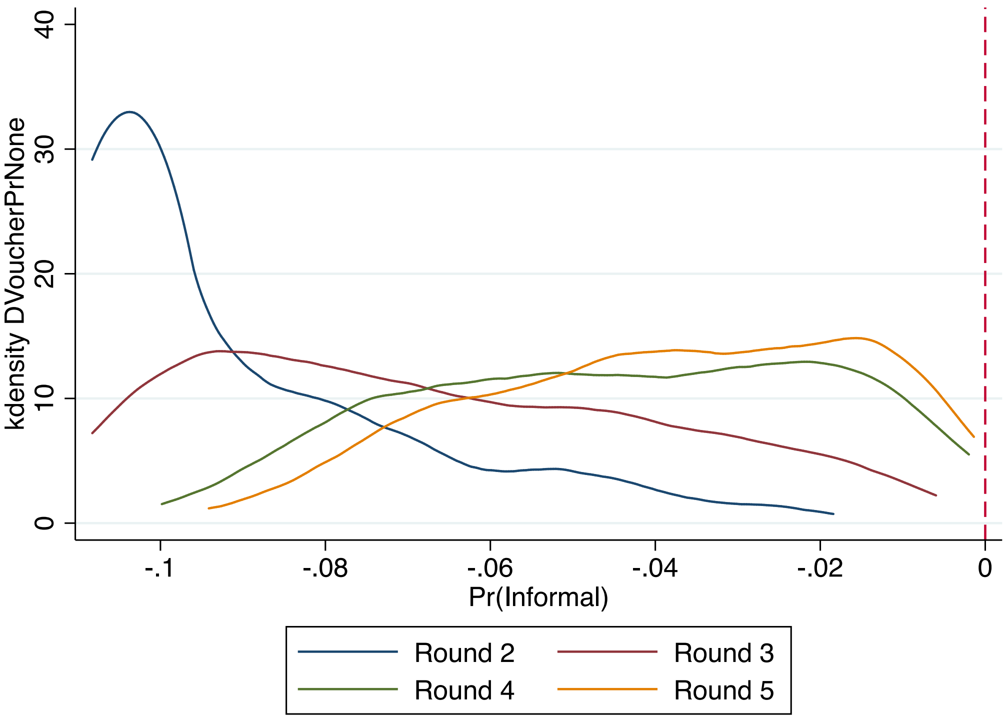
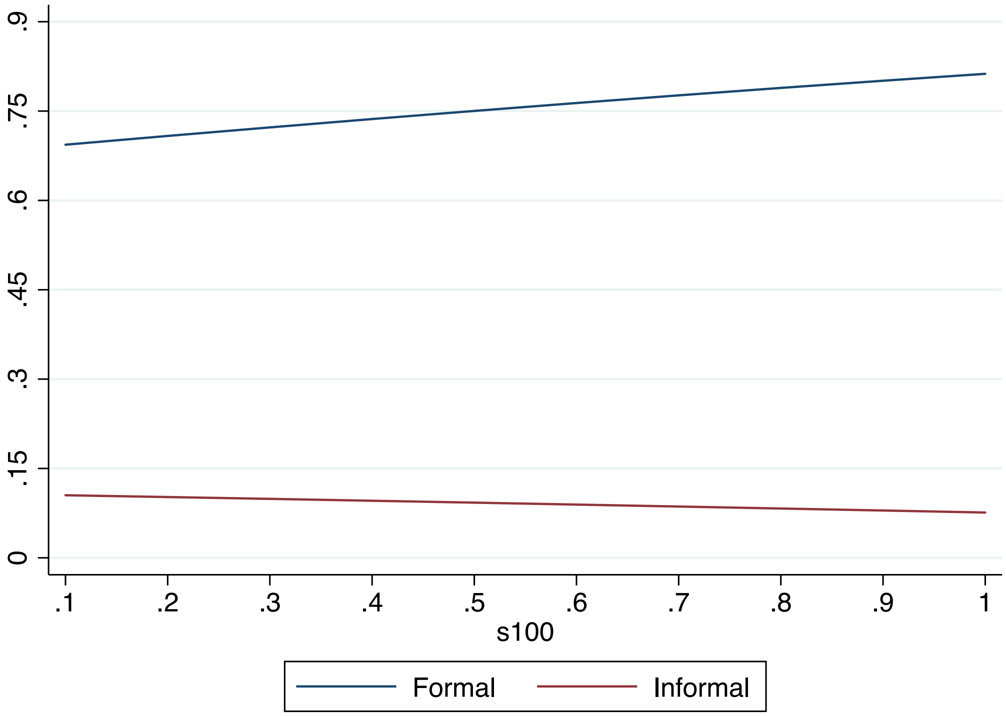


Figure 5. Spillovers of Voucher Allocation (Ordered Probit)

(a) s100 and Connection Type



(b) Marginal Effects of s100

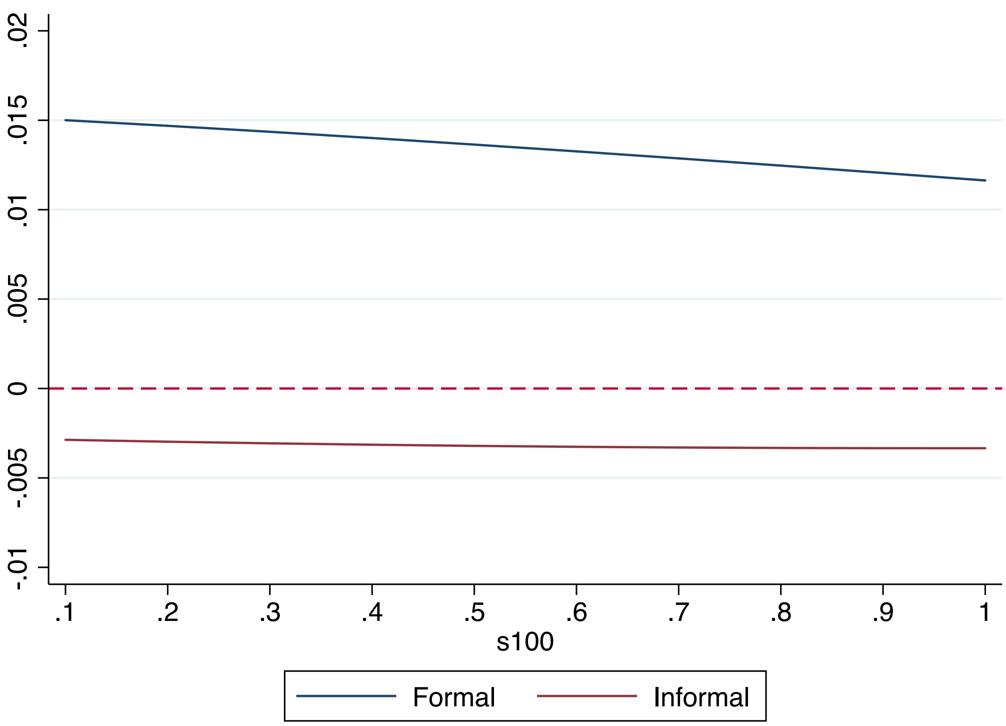


Table 1A - Summary Statistics and Balance by Treatment Arm

	Control Group	20% Discount	Diff: C-20%	50% Discount	Diff: C-50%
Age of household head	49.20	50.80	-1.60	48.99	0.21
	(1.47)	(1.25)	(1.92)	(1.29)	(1.96)
Household head is male	0.62	0.72	-0.10*	0.72	-0.10*
	(0.04)	(0.03)	(0.05)	(0.03)	(0.05)
Household size	4.19	4.65	-0.46*	4.82	-0.63**
	(0.18)	(0.19)	(0.27)	(0.18)	(0.27)
Total dependency ratio	0.47	0.44	0.02	0.43	0.03
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
Maximum schooling in the household	5.51	5.76	-0.26	5.76	-0.26
	(0.33)	(0.33)	(0.47)	(0.32)	(0.47)
Schooling of the household head	1.90	2.03	-0.14	2.23	-0.33
	(0.25)	(0.25)	(0.36)	(0.26)	(0.37)
Household head is literate	0.49	0.49	-0.00	0.52	-0.03
	(0.04)	(0.04)	(0.06)	(0.04)	(0.06)
Income pc, 1000USD per year	0.55	0.52	0.03	0.57	-0.02
	(0.12)	(0.07)	(0.13)	(0.08)	(0.14)
Monthly expenditure in kerosene	2.96	2.56	0.41	2.20	0.76
	(0.39)	(0.32)	(0.50)	(0.27)	(0.46)
Monthly expenditure in propane	1.69	2.11	-0.42	1.78	-0.09
	(0.25)	(0.22)	(0.33)	(0.22)	(0.33)
Monthly expenditure in candles	0.57	0.55	0.01	0.55	0.01
	(0.14)	(0.13)	(0.19)	(0.13)	(0.19)
Monhtly expenditure in car battery rchg	0.12	0.04	0.08	0.12	0.00
	(0.06)	(0.03)	(0.06)	(0.05)	(0.07)
Cooks with wood	0.76	0.73	0.04	0.73	0.03
	(0.04)	(0.03)	(0.05)	(0.03)	(0.05)
Informal electricity	0.39	0.50	-0.11*	0.48	-0.09*
	(0.04)	(0.04)	(0.06)	(0.04)	(0.06)
Agrees with the following statement					
Electricity illuminates better than kerosene.	0.96	0.95	0.01	0.97	-0.00
	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)
Powering a TV is cheaper w/elect than battery.	0.79	0.74	0.05	0.81	-0.03
	(0.04)	(0.03)	(0.05)	(0.03)	(0.05)
Cooking with electricity is not convenient	0.61	0.46	0.15***	0.50	0.11*
	(0.04)	(0.04)	(0.06)	(0.04)	(0.06)
Electricity is very expensive	0.54	0.43	0.10*	0.47	0.06
	(0.04)	(0.04)	(0.06)	(0.04)	(0.06)
Woodsmoke generates respiratory problems	0.87	0.84	0.04	0.87	-0.00
	(0.03)	(0.03)	(0.04)	(0.02)	(0.04)
Kerosene is not an expensive source of lighting	0.42	0.35	0.07	0.32	0.10*
	(0.04)	(0.04)	(0.06)	(0.03)	(0.05)
Kerosene is the best way to illuminate my household	0.28	0.20	0.08	0.20	0.08
	(0.04)	(0.03)	(0.05)	(0.03)	(0.05)

Notes: Columns 1, 2, and 4 show the mean values for each of the treatment arms at baseline (standard errors in parentheses). Column 3 and 5 report the difference in means between the control group and households that received a 20% or 50% discount voucher, respectively (standard errors in parentheses). Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Table 1B - Validating the Randomization of Voucher Density, OLS estimates

	(1) Age Head	(2) Head Male	(3) Head Schooling	(4) Household Size	(5) Adequate Walls	(6) Income	(7) Group Member
Vouchers, 0-100m radius	0.995*	0.004	0.189*	-0.017	-0.007	0.008	0.012
	(0.525)	(0.015)	(0.098)	(0.078)	(0.013)	(0.037)	(0.011)
Vouchers, 100-200m radius	0.589	-0.000	-0.022	-0.072	0.022	0.010	0.005
	(0.565)	(0.017)	(0.106)	(0.083)	(0.014)	(0.039)	(0.012)
Vouchers, 200-300m radius	0.036	0.001	-0.005	-0.043	0.003	0.043	-0.002
	(0.453)	(0.013)	(0.085)	(0.067)	(0.011)	(0.032)	(0.010)
Eligible Neighbors, 0-100m radius	-1.157**	0.002	-0.026	0.015	-0.001	-0.035	-0.005
	(0.496)	(0.015)	(0.093)	(0.073)	(0.012)	(0.035)	(0.011)
Eligible Neighbors, 100-200m radius	-0.452	-0.001	-0.095	-0.026	-0.019**	0.007	0.004
	(0.315)	(0.009)	(0.059)	(0.046)	(0.008)	(0.022)	(0.007)
Eligible Neighbors, 200-300m radius	0.073	0.002	0.027	0.027	-0.028***	-0.025	-0.003
	(0.246)	(0.007)	(0.046)	(0.036)	(0.006)	(0.017)	(0.005)
Mean Dependent Variable	49.76	0.69	2.08	4.58	0.76	0.55	0.13
Number of Observations	486	486	486	486	486	486	486
R squared	0.184	0.054	0.246	0.065	0.199	0.081	0.030

Notes: The dependent variable in each regression is indicated in the column name. Adequate walls indicates adobe, brick, or concrete walls; group member indicates whether any of the household members is a community group member. The controls in each regression include age, sex, and schooling of the household head; number of household members; an indicator for adobe, brick or concrete walls; monthly kerosene expenditure; per capita income; and an indicator for households that have at least one community group member; but when any of these is the dependent variable, it is not included as an explanatory variable. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Table 2 - Discount Vouchers and Connection to the Grid (LPM)

	(1) Connected	(2) Connected	(3) Connected	(4) Connected
Voucher x Post	0.120***	0.120***		
	(0.041)	(0.041)		
Voucher x Round 2			0.134**	0.131**
			(0.054)	(0.053)
Voucher x Round 3			0.147***	0.148***
			(0.054)	(0.054)
Voucher x Round 4			0.105**	0.109**
			(0.046)	(0.046)
Voucher x Round 5			0.095**	0.096**
			(0.042)	(0.043)
s100 x Post	0.130***	0.130***		
	(0.037)	(0.037)		
s100 x Round 2			0.123**	0.126**
			(0.054)	(0.054)
s100 x Round 3			0.156***	0.156***
			(0.049)	(0.049)
s100 x Round 4			0.109***	0.107***
			(0.041)	(0.041)
s100 x Round 5			0.134***	0.132***
			(0.036)	(0.036)
Household Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	2269	2269	2269	2269
Number of Households	494	494	494	494
Mean Dep.Var.	0.56	0.56	0.56	0.56
R-squared	0.45	0.63	0.45	0.63

Notes: The dependent variable in all cases is an indicator of formal connection to the grid. "s100" is the share of eligible neighbors within 100m that received a voucher. "s200" is the share of eligible neighbors between 100-200m radius that received a voucher. Standard errors clustered at the household level, reported in parentheses. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Table 3 - Connection Fee and Connection to the Grid (LPM)

	(1) Connected	(2) Connected	(3) Connected	(4) Connected	(5) Connected
Fee x Post	-0.203**	-0.201**			
	(0.085)	(0.085)			
Fee x t					-0.327***
					(0.122)
Fee x t-squared					0.054***
					(0.020)
Fee x Round 2			-0.324***	-0.314***	
			(0.115)	(0.114)	
Fee x Round 3			-0.260**	-0.261**	
			(0.108)	(0.108)	
Fee x Round 4			-0.126	-0.128	
			(0.094)	(0.094)	
Fee x Round 5			-0.108	-0.106	
			(0.088)	(0.088)	
Fee100 x Post	-0.348***	-0.349***			
	(0.099)	(0.099)			
Fee100 x t					-0.400***
					(0.146)
Fee100 x t-squared					0.055**
					(0.024)
Fee100 x Round 2			-0.294**	-0.303**	
			(0.144)	(0.143)	
Fee100 x Round 3			-0.450***	-0.447***	
			(0.127)	(0.126)	
Fee100 x Round 4			-0.297***	-0.296***	
			(0.108)	(0.108)	
Fee100 x Round 5			-0.352***	-0.349***	
			(0.097)	(0.098)	
Household Fixed Effects	No	Yes	No	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	No
Lin + Quad Time Trend	No	No	No	No	Yes
Observations	2269	2269	2269	2269	2269
Number of Households	494	494	494	494	494
Mean Dep.Var.	0.56	0.56	0.56	0.56	0.56
R-squared	0.44	0.62	0.44	0.63	0.62

Notes: The dependent variable in all cases is an indicator of formal connection to the grid. "fee100" is the average fee for households within 100m. "fee200" is the average fee for neighbors between 100-200m radius. Standard errors clustered at the household level, reported in parentheses. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Table 4 - Connection to the Grid and Externalities (IV)

	(1) Connected	(2) Connected	(3) Connected	(4) Connected	(5) Connected	(6) Connected
Voucher x Post	0.122***	0.124***				
	(0.040)	(0.040)				
Fee x Post			-0.191**	-0.199**		
			(0.082)	(0.084)		
Fee x t					-0.324***	-0.332***
					(0.116)	(0.117)
Fee x t-squared					0.054***	0.054***
					(0.019)	(0.019)
$\overline{E}$ (percentage)	0.193***		0.222***		0.214***	
	(0.055)		(0.058)		(0.058)	
$\overline{E}$ (number)		0.103***		0.111***		0.101***
		(0.030)		(0.031)		(0.030)
Household Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2256	2256	2256	2256	2256	2256
Number of Households	481	481	481	481	481	481
Mean Dep.Var.	0.56	0.56	0.56	0.56	0.56	0.56
R-squared	0.64	0.64	0.64	0.64	0.64	0.64

Notes: The dependent variable in all cases is an indicator of formal connection to the grid. $\overline{E}$ denotes formal electric connections within 100m of the household. Excluded instrument in columns (1) and (2) is s100; in columns (3) and (4), fee100. Columns (2) and (4) control for the total number of eligible neighbors within 100m. Standard errors clustered at the household level, reported in parentheses. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Appendix - Connection to the Grid and Externalities (IV)

	(1) Connected	(2) Connected	(3) Connected	(4) Connected
Fee x t	-0.324***	-0.332***	-0.322***	-0.329***
	(0.116)	(0.117)	(0.116)	(0.116)
Fee x t-squared	0.054***	0.054***	0.054***	0.054***
	(0.019)	(0.019)	(0.019)	(0.019)
$\overline{E}$ (percentage)	0.214***		0.214***	
	(0.058)		(0.058)	
$\overline{E}$ (number)		0.101***		0.100***
		(0.030)		(0.030)
Lin + Quad Time Trend	No	No	Yes	Yes
Household Fixed Effects	Yes	Yes	No	No
Observations	2256	2256	2256	2256
Number of Households	481	481	481	481
Mean Dep.Var.	0.56	0.56	0.56	0.56
R-squared	0.64	0.64	0.64	0.64

Notes: The dependent variable in all cases is an indicator of formal connection to the grid. $\overline{E}$ denotes formal electric connections within 100m of the household. Excluded instruments in all columns are fee100 $\times$ t and fee100 $\times$ t^2. Columns (2) and (4) control for the total number of eligible neighbors within 100m. Standard errors clustered at the household level, reported in parentheses. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Table 5 - Switching (LPM)

	None-Formal		Informal-Formal		None-Informal	
	(1) Switched	(2) Switched	(3) Switched	(4) Switched	(5) Switched	(6) Switched
voucher	0.010	-0.031	0.253***	0.154***	-0.049	0.031
	(0.057)	(0.067)	(0.050)	(0.059)	(0.048)	(0.061)
s100	0.147**	0.103	0.086*	0.111**	-0.009	-0.056
	(0.059)	(0.065)	(0.048)	(0.049)	(0.049)	(0.059)
Baseline Covariates	No	Yes	No	Yes	No	Yes
Subdistrict Fixed Effects	No	Yes	No	Yes	No	Yes
Observations	275	265	264	259	275	265
Mean Dep.Var.	0.75	0.76	0.85	0.85	0.15	0.15
R-squared	0.02	0.29	0.12	0.40	0.00	0.16

Notes: The dependent variable in all cases is an indicator of formal connection to the grid. $\overline{E}$ denotes formal electric connections within 100m of the household. Excluded instrument in columns (1) and (2) is s100; in columns (3) and (4), fee100. Columns (2) and (4) control for the total number of eligible neighbors within 100m. Standard errors clustered at the household level, reported in parentheses. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Table 6 - Multinomial Choice (Ordered Probit)

	(1) Connection Type	(2) Connection Type	(3) Connection Type	(4) Connection Type
type				
voucher	0.248**	0.272**		
	(0.117)	(0.125)		
s100	0.498***	0.501***		
	(0.130)	(0.135)		
Fee			-0.435	-0.547*
			(0.271)	(0.281)
Fee100			-1.188***	-1.261***
			(0.389)	(0.406)
cut1				
Constant	-0.034	-0.273	-1.766***	-2.153***
	(0.096)	(0.330)	(0.382)	(0.511)
cut2				
Constant	0.321***	0.118	-1.416***	-1.766***
	(0.100)	(0.335)	(0.375)	(0.507)
Baseline Covariates	No	Yes	No	Yes
Observations	1767	1714	1767	1714

Notes: The dependent variable in all cases is an indicator of formal connection to the grid. $\overline{E}$ denotes formal electric connections within 100m of the household. Excluded instrument in columns (1) and (2) is fee100; in columns (3) and (4), fee100. Columns (2) and (4) control for the total number of eligible neighbors within 100m. Standard errors clustered at the household level, reported in parentheses. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Table 7 - Characteristics of Adopters

	(1) Voucher	(2) No Voucher
Age of household head	-0.000	0.000
	(0.001)	(0.001)
Gender of household head	-0.036	-0.047
	(0.042)	(0.042)
Literacy of household head	0.011	0.031
	(0.042)	(0.041)
Property title	0.082**	0.101***
	(0.039)	(0.039)
Household size	0.019**	0.028***
	(0.008)	(0.008)
Adequate wall materials	0.026	-0.082
	(0.056)	(0.051)
Dirt floor	-0.104**	-0.082*
	(0.046)	(0.046)
Wood for cooking	-0.012	0.003
	(0.051)	(0.050)
Income (thousand USD)	0.014**	0.008
	(0.007)	(0.007)
Informal connection at baseline	0.177***	0.110**
	(0.050)	(0.045)
Constant	0.686***	0.728***
	(0.114)	(0.112)
Subdistrict FE	No	Yes
Observations	349	349

Notes: Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Appendix - Discount Vouchers and Connection to the Grid (LPM)

	(1) Connected	(2) Connected	(3) Connected	(4) Connected
Voucher x Post	0.115***	0.116***		
	(0.042)	(0.042)		
Voucher x Round 2			0.141***	0.138**
			(0.054)	(0.054)
Voucher x Round 3			0.149***	0.150***
			(0.055)	(0.055)
Voucher x Round 4			0.092*	0.096**
			(0.047)	(0.047)
Voucher x Round 5			0.080*	0.081*
			(0.042)	(0.043)
s100 x Post	0.126***	0.126***		
	(0.037)	(0.037)		
s100 x Round 2			0.132**	0.136**
			(0.055)	(0.054)
s100 x Round 3			0.158***	0.158***
			(0.050)	(0.049)
s100 x Round 4			0.101**	0.099**
			(0.041)	(0.041)
s100 x Round 5			0.122***	0.121***
			(0.036)	(0.036)
s200 x Post	0.027	0.025		
	(0.035)	(0.035)		
s200 x Round 2			-0.039	-0.043
			(0.054)	(0.054)
s200 x Round 3			-0.010	-0.014
			(0.048)	(0.048)
s200 x Round 4			0.068*	0.066*
			(0.038)	(0.038)
s200 x Round 5			0.086***	0.084**
			(0.033)	(0.033)
Household Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	2269	2269	2269	2269
Number of Households	494	494	494	494
Mean Dep.Var.	0.56	0.56	0.56	0.56
R-squared	0.45	0.63	0.45	0.63

Notes: The dependent variable in all cases is an indicator of formal connection to the grid. "s100" is the share of eligible neighbors within 100m that received a voucher. "s200" is the share of eligible neighbors between 100-200m radius that received a voucher. Standard errors clustered at the household level, reported in parentheses. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Appendix- Connection Fee and Connection to the Grid (LPM)

	(1) Connected	(2) Connected	(3) Connected	(4) Connected	(5) Connected
Fee x Post	-0.185** (0.087)	-0.185** (0.087)			
Fee x t					-0.332*** (0.127)
Fee x t-squared					0.056*** (0.020)
Fee x Round 2			-0.330*** (0.117)	-0.323*** (0.116)	
Fee x Round 3			-0.256** (0.111)	-0.259** (0.111)	
Fee x Round 4			-0.091 (0.097)	-0.093 (0.097)	
Fee x Round 5			-0.072 (0.089)	-0.071 (0.089)	
Fee100 x Post	-0.334*** (0.100)	-0.336*** (0.100)			
Fee100 x t					-0.407*** (0.146)
Fee100 x t-squared					0.057** (0.024)
Fee100 x Round 2			-0.301** (0.145)	-0.313** (0.144)	
Fee100 x Round 3			-0.448*** (0.128)	-0.448*** (0.127)	
Fee100 x Round 4			-0.275** (0.108)	-0.274** (0.108)	
Fee100 x Round 5			-0.324*** (0.098)	-0.323*** (0.098)	
Fee200 x Post	-0.115 (0.090)	-0.108 (0.091)			
Fee200 x t					0.034 (0.144)
Fee200 x t-squared					-0.018 (0.024)
Fee200 x Round 2			0.032 (0.137)	0.046 (0.137)	
Fee200 x Round 3			-0.030 (0.126)	-0.019 (0.126)	
Fee200 x Round 4			-0.216** (0.096)	-0.214** (0.096)	
Fee200 x Round 5			-0.327***	-0.323***	

Table 5 - Switching (LPM)

	None-Formal		Informal-Formal		None-Informal	
	(1)	(2)	(3)	(4)	(5)	(6)
	Switched	Switched	Switched	Switched	Switched	Switched
Fee	0.055	0.076	-0.365***	-0.122	0.085	-0.056
	(0.127)	(0.138)	(0.106)	(0.113)	(0.105)	(0.125)
Fee100	-0.261*	-0.157	-0.392***	-0.360***	0.021	0.088
	(0.155)	(0.167)	(0.131)	(0.132)	(0.128)	(0.152)
Baseline Covariates	No	Yes	No	Yes	No	Yes
Subdistrict Fixed Effects	No	Yes	No	Yes	No	Yes
Observations	275	265	264	259	275	265
Mean Dep.Var.	0.75	0.76	0.85	0.85	0.15	0.15
R-squared	0.01	0.29	0.08	0.39	0.00	0.16

Notes: The dependent variable in all cases is an indicator of formal connection to the grid. $\overline{E}$ denotes formal electric connections within 100m of the household. Excluded instrument in columns (1) and (2) is fee100; in columns (3) and (4), fee100. Columns (2) and (4) control for the total number of eligible neighbors within 100m. Standard errors clustered at the household level, reported in parentheses. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Appendix - Multinomial Choice Round by Round (Ordered Probit)

	Round 2		Round 3		Round 4		Round 5	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
main								
voucher	0.211		0.304**		0.299*		0.309*	
	(0.136)		(0.147)		(0.161)		(0.173)	
s100	0.330**		0.538***		0.521***		0.763***	
	(0.142)		(0.161)		(0.183)		(0.204)	
Fee		-0.559*		-0.623*		-0.442		-0.530
		(0.294)		(0.328)		(0.360)		(0.386)
Fee100		-0.750**		-1.477***		-1.341***		-1.865***
		(0.374)		(0.438)		(0.486)		(0.549)
cut1								
Constant	-0.557	-1.907***	-0.729*	-2.887***	-1.176***	-3.050***	-1.062**	-3.553***
	(0.344)	(0.508)	(0.388)	(0.586)	(0.424)	(0.638)	(0.447)	(0.710)
cut2								
Constant	-0.002	-1.354***	-0.411	-2.572***	-0.851**	-2.729***	-0.787*	-3.285***
	(0.343)	(0.506)	(0.387)	(0.584)	(0.422)	(0.636)	(0.446)	(0.707)
Observations	414	414	422	422	446	446	432	432

Notes: The dependent variable in all cases is an indicator of formal connection to the grid. $\overline{E}$ denotes formal electric connections within 100m of the household. Excluded instrument in columns (1) and (2) is fee100; in columns (3) and (4), fee100. Columns (2) and (4) control for the total number of eligible neighbors within 100m. Standard errors clustered at the household level, reported in parentheses. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*