Household Electri cation and Indoor Air Pollution

Manuel Barron and Maximo Torero

University of Berkely, IFPRI

January 2015

Household Electrification and Indoor Air Pollution

Manuel Barron[1] and Maximo Torero[2]

This Version: January 16th, 2015
(The latest version is available here)[3]

Abstract

This paper provides the first empirical evidence that household electrification leads to direct and substantial welfare improvements via reductions in indoor air pollution. In the setting of a recent electrification program in northern El Salvador, we exploit a unique dataset on minute-by-minute fine particulate matter ($PM_{2.5}$) concentration within the framework of a clean experimental design. Two years after baseline, overnight $PM_{2.5}$ concentration was on average 67% lower among households that were randomly encouraged to connect compared to those that were not. This change is driven by reductions in kerosene use. As a result, the incidence of acute respiratory infections among children under 6 fell by 65% among connected households. Estimates of exposure measures suggest large health gains for all household members, but these gains are unequally distributed by gender. In addition, we show that when the electrification rate among the non-encouraged group caught up with that of the encouraged group, the effects in the former group were similar to those in the latter.

JEL codes: Q53, Q56, I15, I18, O13, O33

keywords: household electrification, indoor air pollution, health

---

[1]Department of Economics, UC Santa Cruz. mbarron4@ucsc.edu. Latest version: http://people.ucsc.edu/~mbarron4/research.html

[2]Markets, Trade, and Institutions Division, IFPRI. m.torero@cgiar.org

1   Introduction

In 2009, 1.3 billion people lacked access to electricity at home (International Energy Agency, 2011). At night, households with no access to electricity make do mostly with candles or kerosene lamps to satisfy their illumination needs. These sources of light provide poor illumination and, more importantly, emit high amounts of pollutants harmful for human health. In fact, indoor air pollution (IAP) is the third leading risk factor for global disease burden, after high blood pressure and smoking (Lim et al., 2013).[1] Given the stylized fact that lighting is one of the first uses of electricity in newly electrified areas (see e.g. Bernard, 2012; Barnes, 2007; Independent Evaluation Group, 2008), electrification is expected to decrease IAP levels by replacing traditional source of lighting, like kerosene, candles, and woodsticks. These reductions and their potential health effects are often argued to be one of the main benefits of electrification, but there is no solid empirical evidence to date.

Our paper contributes to filling this gap by providing the first experimental estimates of the relationship between household electrification and indoor air pollution. To answer this key question we collected, within the frame of a clean experimental design, a uniquely rich dataset that pairs minute-by-minute fine particulate matter ($PM_{2.5}$) concentration with detailed data on household members' time allocation. This unique combination of data and experimental framework allows an accurate estimation of the lower bound of the changes in $PM_{2.5}$ concentration driven by electrification and, moreover, to assess the magnitude of the health effects that these changes imply.

Our research question is situated within the broader area of the effects of electrification, an active area of research in which the debate is far from settled. The massive amounts of resources allocated to rural electrification[2] have usually been justified on the assumed benefits of electrification on health, education, and income, but most of the empirical evidence on which these claims are based is weak (Bernard, 2012; Independent Evaluation Group, 2008), and the more recent literature shows mixed results. Some of the recent evidence suggests that electricity may save time in household chores, thus increasing female labor supply (Dinkelman, 2011; Grogan and Sadanand, 2012) or that it leads to improvements in educational outcomes, consumption, and income (e.g. Khandker, Barnes, and Samad; van de Walle et al., 2013), while others find no such relationships (e.g. Bernard and Torero, 2013;

---

[1]Ambient pollution also has important negative health effects, as shown by Chay and Greenstone (2003) and more recently by, e.g., Chen et al. (2013) and Hanna and Oliva (2011). This paper will concentrate on indoor air pollution only.

[2]For instance, the World Bank recommends investing $10 billion per year between 2010 and 2020 to production and distribution of electricity in rural areas in Africa (World Bank, 2009). Given that the institution aims to provide 250 million people across Africa with modern sources of energy by 2030 (World Bank, 2007), understanding the effects of electrification is of urgent importance.

Bensch, Kluve, and Peters, 2011).

Our estimates provide the first experimental evidence that household electrification leads to substantial, immediate and sustained improvements on welfare by improving indoor air quality. Overnight $PM_{2.5}$ concentration was 67 percent lower among voucher recipients than among non-recipients one year after the electricity rollout. Due to imperfect compliance with voucher assignment, this is a lower bound of the true effect of electrification on $PM_{2.5}$ concentration. Similarly, the probability probability of observing concentrations above 67 $\mu g/m^3$ between 5pm and 7am fell by at least 33 percentage points. Although there are no clear benchmarks in the environmental health literature, we consider this a useful reference point because over a 24-hour period, a concentration of 67 $\mu g/m^3$ implies the same $PM_{2.5}$ exposure as smoking one cigarette (1200 $\mu g/m^3$).[3]

The reductions in overnight $PM_{2.5}$ concentration result in large and significant falls in acute respiratory infections (ARI) among children under 6. Depending on the exact specification voucher recipients report 18 percentage points lower incidence of ARI in the four weeks preceding the survey than non-recipients, translated to a 65 percentage point reduction (p¡0.10) as a consequence of electrification over the course of the study. To assess further health implications of the observed reductions in $PM_{2.5}$ concentration among the population over 6, in section 6.1 we use data from the time allocation module to estimate the change in daily exposure to $PM_{2.5}$. The resulting reductions in exposure to $PM_{2.5}$ are large but unequally distributed among household members. Adult males benefit the most, with 59% lower exposure. Since adult females are still exposed to high $PM_{2.5}$ concentrations while cooking, they benefit the least, with reductions in exposure of 33%. The figures for children are 46% (males) and 39% (females). The dose-response function recently developed by Pope III et al. (2011) based on first and second-hand tobacco smoking associates the figures we find with large reductions in the relative risk of lung cancer, 25% for adult females, 33% adult males, with the respective figures for children falling 25% (females) and from 30% (males). Although the composition of $PM_{2.5}$ generated by kerosene combustion is not the same as the one generated by cigarette smoking, the current scientific evidence cannot reject that their health effects are similar.

The main mechanism behind the $PM_{2.5}$ reductions in our study setting is a substitution away from kerosene lighting. Electrification caused large reductions in kerosene expenditures, while changes in other traditional lighting sources like candles are small in magnitude and not statistically significant. We find no evidence of changes in cooking practices either. The reduction in kerosene use has important health implications, because although kerosene is

---

[3]Concentration is the amount of $PM_{2.5}$ per unit of space, usually measured in mg/m^3 or $\mu g/m^3$. Exposure is the amount of $PM_{2.5}$ that effectively enters the human body, usually measured in mg or μg.

usually considered a cleaner alternative to biomass, emissions from kerosene-burning devices are considered extremely harmful for human health. Aside from $PM_{2.5}$, kerosene emissions include carbon monoxide (CO), nitric oxides (NOX), and sulfur dioxide (SO2) (Schare and Smith, 1995; Fan and Zhang, 2001; Röllin et al., 2004). These pollutants can impair lung function and increase infectious illness, asthma, and cancer risks Lam et al. (2012a). The reduction in kerosene use also has important environmental consequences, since kerosene lighting (let alone other uses like cooking or heating) is responsible for 3 percent of annual black carbon emissions globally (Lam et al., 2012b).

[REWRITE THIS PARAGRAPH] In a companion paper (Barron and Torero, 2013a) we study the changes in time allocation caused by household electrification, uncovering some interesting relationships. First, electrification causes school-age (6 to 15 year old) children to study one additional hour per week. It also has an effects on the extensive margin, increasing the probability of studying by 17 percentage points.[4] Adult males spend roughly an additional hour in farm work daily, while reducing leisure by a similar amount. This is consistent with van de Walle et al. (2013)'s argument that people require some amount of "quality-time" leisure per day. Without electrification it is not possible to have quality-time leisure in the evenings, thus adult males quit their job early to ensure enough leisure in the afternoon. With electrification, they can enjoy quality time leisure at night, thus it is now possible to spend more time at work and still enjoy leisure at night. Third, adult women are 50% more likely to participate in market activities (selling goods), from a baseline value of 3 percentage points. Although large with respect to the baseline value, it is a smaller labor supply effect than the one found by Dinkelman (2011).

The reductions in $PM_{2.5}$ concentration found in this study are not necessarily obvious ex-ante for four reasons. First, since households continue to use fuelwood for cooking, and woodsmoke produces higher IAP concentration than lighting fuels, the resulting reductions in IAP may have been small and thus irrelevant for policy. In fact, given that there is still scarce evidence in the specialized literature, the evidence of a strong, positive relationship between kerosene use and $PM_{2.5}$ concentration in a sample in which 70% of households rely on fuelwood for cooking is a contribution to the environmental health literature.[5] Second, it may be the case that only the heaviest kerosene users experience significant reductions, and thus the average reduction may just be an illusion. We explore this issue and find even

---

[4]Since this time is being spent in an environment with lower pollutant concentration, in a smokeless, better illuminated room, each minute of studying is more productive as well. Adding the health effects of the reductions in exposure to pollutants should increase the productivity of study time even further. Hence we argue that there are gains in learning not only due to extended study time, but also in quality of study time. Whether this will reflect in more years of schooling is another question, but it should unequivocally translate into better learning.

[5]We explore this issue in closer detail in Barron and Torero (2013b).

though the highest polluters indeed get the highest gains, 80% of households experienced significant reductions in overnight $PM_{2.5}$ concentration. Third, voucher recipients may have not adopted electricity at a high enough rate, thus the relationship between receiving a voucher and connecting to the grid may be too weak to reflect in IAP measurements. Fourth, households may be effective in dealing with IAP from lighting sources.[6]

Our findings differ qualitatively from the typical findings in the literature on improved cookstoves in two dimensions. First, the effects found in field studies on improved cookstoves are null or small, especially when compared to the effects expected from laboratory or controlled field studies.[7,8] The second dimension is that unlike the cookstove literature (e.g. Hanna, Duflo, and Greenstone, 2012), the reductions in $PM_{2.5}$ observed in our sample are steady over time. The time resilience of the effects we find strengthens the link between household electrification and human health discussed in the preceding paragraphs.[10]

The remainder of this document is organized as follows. Section 2 presents the study setting and discusses the data. Section 3 describes the conceptual framework that guides our study. Section 4 presents the econometric approach. We discuss the main results in section 5. In section 6 we combine the findings on $PM_{2.5}$ concentration with time-use data to infer $PM_{2.5}$ exposure and health implications. Section 7 presents a note on the profitability of replicating our by a private agent, and section 8 concludes.

2 The Study Setting

In this section we describe the study setting and give detailed information on the experimental design. After a brief introduction on fine particulate matter and a description of the channels through which it affects human health, we describe the protocol and equipment used to measure it. Next we present descriptive statistics of our baseline survey and show that groups resulting from the randomization have balanced means across a wide set of observable characteristics.

---

[6]In fact, households in the study setting households usually place the kerosene lantern or *ocote* sticks close to windows and doors. Ocote is a type of wood used for ligthing by the poorest households.

[7]By "controlled field studies" we refer to field trials in which field officers or research assistants constantly visit households to make sure the improved cookstoves are being used the way they are supposed to be used. Although this blocks the behavioral response of the household, the evidence is informative as to the maximum potential benefits of new technologies under adequate use.

[8]The main driver of this divergence is household behavior: households do not adopt the improved stoves (WHO, 2006; Mobarak et al., 2012)[9], or if they do adopt them, do not use them correctly (Wallmo, Jacobson et al., 1998). A large body of literature investigates the reasons (e.g. Masera, Díaz, and Berrueta, 2005; Pokharel, 2003; Troncoso et al., 2007).

[10]Although recent work suggests that exposure to IAP for short periods of time can have permanent negative effects on in-utero babies.

2.1 The Electrification Program and Experimental Design

The study takes place during a recent grid extension and intensification program in northern El Salvador, designed to be rolled-out in three phases according to construction costs and accessibility. In this program, the El Salvadorian government covered all the installation costs up to the electric meter, and households had to pay for their internal wiring and a connection fee (for a safety certification). The fee for the safety certification is of around US$ 100. It is non-trivial for a household, amounting to roughly 20% of annual per capita income in our sample.

The experimental sample consists of 500 households located in subdistricts that were scheduled to be covered by the program during its first year. We generated experimental variation in the connection fee by offering discount vouchers to a randomly selected subsample. We randomly allocated 200 low-discount vouchers (20% discount), 200 high-discount vouchers (50% discount), and left the remainder households as control group (N=100). The exogenous variation in the connection fee generated by the random voucher allocation deals with self-selection in connection to the grid. Vouchers were valid for a discount towards the safety certification to be reimbursed after paying the full cost. Each voucher showed the name and address of the beneficiary, it was non-transferable, and it was valid for 9 months.[11]

The random voucher allocation also creates exogenous variation in the number of voucher recipients in a given neighborhood of household i (controlling for the number of eligible neighbors), which generates variation in the number of new connections around household i, so we can control for the role of spillovers on grid connection. The sign of the effect is theoretically ambiguous. On the one hand, observing their neighbors connect to the grid may make households more prone to connect, through a combination of social learning and imitation effects.[12] On the other, higher formal connection rates in a neighborhood reduce the cost of getting an informal connection, so the number of vouchers around a household may increase the number of informal connections. To estimate the role of spillovers on adoption, we use the number of household i's neighbors that received a voucher in a given radius (0-100 meters, 100-200 meters, 200-300 meters), controlling by the number of eligible neighbors in that radius. Eligible households are households with no electricity at baseline.

EHEIPCER, the household survey implemented for this study, is a fairly standard survey that collected data on demographic characteristics, health, education, housing characteris-

---

[11]In a few cases there was a delay in the implementation of the program, so the expiration date was extended for the households from those areas.

[12]Social learning would occur if households observed the private benefits of electrification (better illumination, less smoke at night, better food availability, more enjoyable leisure time) from their neighbors. Imitation effects (also known as "preferences interactions" in the literature) are similar to a "keeping-up with the Joneses" story: a household wants electricity because its neighbors have it.

tics, energy use, income, consumption, among others. In particular, it includes a detailed module on time allocation for up to four household members: the male head, the female head, and up to two school-age children. Strict training sessions were conducted to ensure high quality in data collection, which was conducted with handheld computers. Enumerators were trained and selected by the authors with the assistance of *DIGESTYC* (the Salvadorian Bureau of Statistics) and *IFPRI* staff. The indoor air pollution data described below were collected by a subset of enumerators that underwent additional special training to this end.

The baseline household survey, designed using the 2007 Population Census as the sampling framework, was collected in November and December 2009. It covered 4,800 households all over northern El Salvador. Three follow-up surveys have been collected in the same months in 2010, 2011, and 2012 respectively. An additional follow-up survey is scheduled for November 2013, and a final round is scheduled to be fielded in November 2015.

2.2   Fine Particulate Matter (PM$_{2.5}$)

The main outcome of interest in this paper is fine particulate matter (PM$_{2.5}$), which is particulate matter (PM) with a diameter of 2.5 microns or less (1 micron = 0.001 millimeters). Particulate matter, also known as particle pollution, is a complex mixture of small particles and liquid droplets composed by potentially hundreds of chemicals. Given the complexity of their composition, particles are mainly classified according to their size. Particles with a diameter between 10 and 2.5 microns are also known as "coarse particulate matter" or "inhalable coarse particulate matter" (PM$_{10}$), while particulates with a diameter under 2.5 microns are known as "fine particulate matter" (PM$_{2.5}$). As reference, a human hair has a diameter of between 50 to 70 microns, 20 to 30 times larger than the cutoff point for fine particulate matter.

Particle size is inversely linked to its potential for causing health problems. Both PM$_{10}$ and PM$_{2.5}$ can pass through the throat and nose and enter the lungs, but being smaller, PM$_{2.5}$ can get deeper into the lungs and also enter the bloodstream, thus causing more damages to health than PM$_{10}$. Both PM$_{10}$ and PM$_{2.5}$ have been shown to cause or aggravate heart and lung diseases. Further, there is evidence that they weaken the immune system, making the body more vulnerable to disease in general. Thus, it affects cognitive ability and productivity.[13]

---

[13]Other than by size, particles that compose particulate matter are usually classified as primary or secondary. Primary particles are emitted directly by a source, like kerosene lamps, cookstoves, unpaved roads, or construction sites. The outcome of interest in this paper falls in this category. Secondary particles, on the other hand, are formed in the atmosphere, as a result of sulfur dioxides and nitrogen oxides emitted from power plants, industries and automobiles. Secondary particles account for most of the particulate matter in developed countries, while the converse is arguably true in developing regions.

2.3 PM$_{2.5}$ Measurement

A central part of the project consisted in collecting data on overnight PM$_{2.5}$ concentration. We obtained PM$_{2.5}$ measurements in two subsamples of households, one experimental and one non-experimental. The experimental subsample includes PM$_{2.5}$ data on 141 randomly selected households from the 500 households that were considered for voucher allocation. The reasons for not selecting the whole sample were logistical and budgetary. Measurements for these households were collected with rounds 3 and 4 of the household survey. The non-experimental subsample is formed by 200 households of *EHEIPCER* households from neighboring subdistricts in the same departments as the experimental sample (San Miguel and Chalatenango) that had not connected to the grid by September 2010. Measurements in these households collected with rounds 2, 3, and 4 of the household survey. The non-experimental sample is formed by households that had not connected to the grid by round 2. Descriptive statistics of both subsamples are reported in Appendix Table A1.

In each household we measured minute-by-minute PM$_{2.5}$ concentration between 1700 hours and 0700 hours the next morning in the main evening living area. The main evening living area is defined as the room where household members spent most of their time awake during the evenings. In the majority of cases this was the living room. Measurements were conducted with the University of California at Berkeley Particle and Temperature Sensor (UCB-PATS). The UCB-PATS is a small, portable and non-intrusive datalogging particle monitor for indoor environments. It uses a photoelectric detector to measure PM$_{2.5}$ concentrations down to 25 $\mu g/m^3$. The UCB-PATS records PM$_{2.5}$ concentration, relative humidity and temperature at a 1 min time resolution. For details on the development and performance of the UCB-PATS see Litton et al. (2004), Edwards et al. (2006), and Chowdhury et al. (2007).

Experienced and meticulously trained enumerators visited the selected households, explained the purpose of the study and obtained consent to place the UCB-PATS in the home. The protocol implemented to measure PM$_{2.5}$ concentration is similar on the protocol applied by Northcross et al. (2010) for cookstoves. It is a standard protocol in the cookstove literature but there is no standard protocol in place for measuring indoor air pollution emitted by kerosene lanterns.The monitor was placed in the room where most household members spent most of their time awake during the evenings. In practice, this was mostly the living room. In a handful of cases it was the master bedroom. The monitor was placed on a wall 1m (horizontally) from the place where the lamp is usually located in the evenings, at least 1.50m away from any working doors or windows, and at a height of 1.50m above the ground. In the baseline measurement enumerators took pictures of the pictures of the placement to ensure placing the monitor in the same place in the follow-up visits. This reduces the risk

of generating artificial variation in $PM_{2.5}$ concentrations by placing the monitor in different locations.

In follow-up measurements, the enumerators used pictures from previous rounds to place the monitors in the same place as the baseline measurement. The enumerators filled a data sheet with exact details on the height, distance, set-up time, pick-up time, among other information. The monitors were placed in the homes before 1600 hrs. If the monitor was placed in a home between Monday and Thursday, it was picked up the next morning starting around 0800 hrs. If it was placed in a home on a Friday, it was picked up the coming Monday starting around 0800 hrs. This procedure was followed to comply with labor regulations in the government sector. In a subsample of households the measurement took place between 5pm on a Friday and 7am the coming Monday. Following the standard practice in the environmental health literature, the resulting $PM_{2.5}$ concentration for those households was averaged across the three days.

2.4   Data, Descriptive Statistics and Balance

According to the 2007 National Census around 80% of the El Salvadorian population had access to electricity. Although this figure is high, there are strong correlations between socioeconomic status, electrification, and use of traditional fuels for lighting or cooking. Figure 1 shows that the poorest municipalities are the ones with the lowest electrification rates and the highest use of traditional fuels for cooking and lighting.

To illustrate the relationship between kerosene use and indoor air quality in our study setting, Figure 2 shows a non-parametric regression of overnight $PM_{2.5}$ as a function of monthly kerosene expenditure (more details on the variables and the samples in the next section). There is a clear positive relationship between these two variables, suggesting that reductions in kerosene use could generate important improvements in indoor air quality. Kerosene provides an important source of variation in $PM_{2.5}$ even with 70% of households using wood for cooking.

Table 1 shows descriptive statistics split by treatment arm. Column 1 shows the means for the control group, column 2 shows the means for the households that received a 20% discount, and column 4 shows the means for households that received a 50% discount. Columns 3 and 5 test for differences between each of the treatment arms and the control group. Household heads are on average 50 years old, 69% of them are male and have 2.4 years of schooling on average. Literacy rates among household heads are low, with only 54% of them reporting being literate. The average age in the households is 30.8 and households are composed by 4.5 members, with a total dependency ratio of roughly 0.45. Annual income is around US$770

8

per head, roughly US$2.11 per person per day.

The main source of energy expenditure is kerosene (US$2.11 per month) mainly used for lighting, and propane (US$2.09 per month), mainly used for cooking, followed by candles (US$0.46/month) and car battery recharging (US$0.08/month), used to power TV sets. Use of wood for cooking was reported by 70% of households. Thirty-eight percent of households had informal access to electricity at baseline. Informal connections consist on a series of extension cables connected to each other and plugged into a neighbor's sockets. They are at most enough for two lightbulbs and some times a television set. For our purposes, households with informal connections were treated as off-grid. This can attenuate the effects of electrification on indoor air quality if we think that households with informal connections rely less on kerosene for lighting than those with no connection at all. However, since it is difficult for the government or the electric utility to determine if a household has informal access to the grid, we argue that the results from our strategy are more relevant for policy purposes.

Households were also balanced regarding their ex-ante perceptions towards energy sources. The vast majority agreed that electricity illuminates better than kerosene (96%) and that woodsmoke generates respiratory problems (87%). Between 30 to 40% of respondents said that kerosene is not an expensive source of lighting, and 20-30% said it as the best way to illuminate their household.

3   Empirical Approach

In this section we describe the econometric approach on which our empirical estimates are based. First we describe the standard parametric approach used for our main results, both for the experimental and the non-experimental sample. In the appendix we describe Changes-in-Changes, a non-parametric approach that allows estimating the distribution of treatment effects and thus analyzing potential treatment effect heterogeneity.

3.1   First Stage: Voucher Allocation and Electrification

In this section we present the first-stage relationship between voucher allocation and household decision to adopt electricity. Since the effect of the voucher is expected to change in time (being stronger at the beginning and wearing off as, on the one hand, vouchers expire and, on the other, the control group catches up) we estimate an adoption regression per round. For each $t = 2, 3, 4$, we estimate:

$$\text{connected}_{it} = \beta \text{voucher}_i + \gamma_{100} s_i^{0-100} + \omega' X_{i0} + \varepsilon_{it} \tag{1}$$

where connected_{it} indicates that household i had a formal connection at time t, voucher_i indicates whether the household received a discount voucher. To control for spillovers we include s_i^{0-100}, the share of eligible neighbors that received a voucher within a 100 meter radius of household i, X_{i0} are individual baseline characteristics.

In some specifications we estimate an equivalent regression stacking all the rounds together, and interact each coefficient with a round fixed effect, which allows us to estimate a fixed effects panel data model.

$$\text{connected}_{it} = \sum_{t=2}^{4} \beta_t \text{voucher}_i + \sum_{t=2}^{4} \gamma_t^{100} s_i^{0-100} + \mu_i + \nu_t + \varepsilon_{it} \tag{2}$$

, where the variables denote the same as in (1) but in addition μ_i and ν_t denote individual and year fixed effects. Given that the vouchers are randomly allocated, both approaches give the same point estimates. There are slight differences in the standard errors. The main conceptual advantage of (2) over (1) is that it allows testing hypotheses on the coefficients by rounds, although given our limited sample size our study is not powered to perform these tests.

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 (1) and (2).

3.2 Second Stage: Electrification and PM$_{2.5}$ concentration

To study the effects of electricity, we exploit the experimental variation in connection fee described in the preceding section to instrument for connection to the grid. At each round, the first stage is given by (1), and the second stage is given by:

$$y_{ithsm} = \delta \widehat{conn}_{it} + \psi' X_{i0} + \text{hour}_h + \text{subdistrict}_s + \text{monitor}_m + \varepsilon_{it} \tag{3}$$

, where y_{it} indicates the outcome of interest (measures of PM$_{2.5}$), X_{i0} includes baseline covariates, as well as subdistrict, hour of the day and monitor fixed effects, while ε_{it} is a disturbance term. δ is the main coefficient of interest, as it gives the causal effect of connection on the outcome for the population of compliers. The complier subpopulation may be small because either (i) there is small take-up among the encouraged group, or (ii) there is large take-up among the non-encouraged group. In our case, the small complier subpopulation is due to large take-up in the non-encouraged group, especially in rounds 4 and 5.

3.3 Reduced Form: Voucher Allocation and PM$_{2.5}$ concentration

Due to a small complier subpopulation, the IV the point estimates are noisy, which generates large standard errors. To avoid relying in noisy estimates, our main results will be based in the reduced form estimates. Effectively, these estimates report the effect of receiving a voucher on the outcome of interest, and as such, these estimates are informative and valuable from a policy perspective. In addition, note that given imperfect compliance, these estimates represent a lower bound of the effects of electrification.

The reduced form is given by:

$$y_{ishm} = \beta \text{voucher}_i + \omega' X_{i0} + \text{hour}_h + \text{subdistrict}_s + \text{monitor}_m + \varepsilon_i \tag{4}$$

We include hour-of-the-day, subdistrict and subdistrict fixed effects. In addition, we control for PM$_{2.5}$ monitor fixed effects, to control for potential measurement error in the equipment.

We estimate two variants of (4). First, we will exploit the minute-by-minute nature of the data. To allow arbitrary structure in the covariance matrix within a household, we cluster the standard errors at the household level. Second, we collapse the data at the household level and run the regressions at the (in the spirit of a "between" estimator) to show that the significance in the coefficients on the voucher variables is not driven by the large number of

observations, we also collapse the data.[14] Given that the first-stage analysis shows a positive and significant relationship between voucher allocation and grid connection, we find it difficult to argue against electrification being the channel through which vouchers affect IAP.

3.4 Non-experimental Sample

In addition to the first-stage and reduced-form analysis, we present non-experimental estimates as further supportive evidence that the channel through which voucher allocation causally affects $PM_{2.5}$ concentration is in fact household electrification. For this purpose we use a sample of approximately 250 *EHEIPCER* households that had not connected to the grid by round 2. The first $PM_{2.5}$ measurement was conducted in round 2, and follow-up measurements were taken in rounds 3 and 4. Appendix 1 shows some summary statistics describing the sample.

$$y_{it} = \delta \times \text{connected}_{it} + \lambda_t + \mu_i + \varepsilon_{it} \tag{5}$$

, where y_{it} is the outcome of interest for household i at time t, connected_{it} indicates whether the household had a connection to the grid, λ_t capture round fixed effects, and μ_i capture household fixed effects. Causal identification of the parameter of interest δ in this setting requires assuming that connected_{it} is uncorrelated with the disturbance term ε_{it} after controlling for time invariant characteristics. This can never be tested, but one way to show supportive, albeit indirect, evidence on this assumption is showing that the outcome variable followed the same time trend between connected and off-grid households before the former group got a connection.

To test for parallel pre-electrification trends, we use the fact that we observe the households in the non-experimental sample in four points in time, from 2009 to 2013 (although there are no $PM_{2.5}$ measurements in 2009). By 2010 no household in this sample had a grid connection. We define three groups based on the timing of grid connection: T1, which is formed by households that connected by 2011; T2, formed by households that connected by 2012; and C, formed by households that did not connect during the study period. Under the assumption of parallel pre-treatment trends in $PM_{2.5}$ concentration, the trend in T2 should be parallel to that of C between 2010 and 2011.

---

[14]The difference between this procedure and the "between" estimator is that the latter requires collapsing the minute-by-minute data from all the rounds to a single point for each household and estimating the effect based on the cross-sectional variation in the resulting sample.

12

$$y_{it} = \theta_2 \times T2_i + \lambda_t + \mu_i + \varepsilon_{it} \tag{6}$$

In the regressions that test parallel pre-electrification trends in $PM_{2.5}$ we compare T2 to C from 2009 to 2010. We also compare it by hour of the day (from 5pm to 6pm, from 6pm to 7pm, up to 6am to 7am) and fail to reject the hypothesis of parallel pre-electrification trends (i.e., in every specification θ_2 results non-significant).

Furthermore, since these are *EHEIPCER* households, we also have their full survey data starting from 2009. This allows testing for parallel pre-electrification trends in both T1 and T2 compared to C in energy expenditure by source (kerosene, candles, gas) and in use of wood for cooking between 2009 and 2010.

$$y_{it} = \theta_1 \times T1_i + \theta_2 \times T2_i + \lambda_t + \mu_i + \varepsilon_{it} \tag{7}$$

Neither θ_1 nor θ_2 result significant in any specification, thus we fail to reject the hypothesis of parallel pre-electrification trends in every case.

4   Results

This section presents the main results of the paper, starting with the experimental estimates on the reduction in $PM_{2.5}$ concentration. Next, we show some non-experimental estimates to support the fact that the changes are in fact driven by electrification and are not an artifact of the data.

4.1 $PM_{2.5}$ Experimental Estimates

Table 2 reports the first stage results. Discount vouchers increase the probability of adoption of a formal connection. In Barron and Torero (2013a) we find that voucher value had very little effect on take-up, so for the sake of parsimony, in what follows we lump both groups in one variable. Individual discount vouchers made households 8-15 percentage points more likely to connect to the grid. The point estimates are larger in the rounds 2 and 3 (13-15 percentage points), and smaller in the last two rounds (8-11 percentage points), as households from the control group got connected to the grid. Figure 3, based on the estimates of Table 2, depicts the evolution of the electrification rate in both groups, showing how the not encouraged group appears to catch up with the encouraged group towards the last two survey rounds.

There are important spillover effects in rounds 2 and 3, which are captured by *s100*. *s100*

is the share of eligible neighbors within 100 meters that received a voucher. Alternatively, we have included the number of vouchers allocated within 100 meters controlling for the number of eligible neighbors in that radius, finding practically the same results. An increase of 10 percentage points in *s100* increases the probability of connection by 1.2-1.6 percentage points. We analyze these effects elsewhere (Barron and Torero, 2013a).

Column 2 includes household fixed effects. The fact that the coefficients are almost exactly equal between columns 1 and 2 reassures that *voucher* and *s100* are uncorrelated with the disturbance terms, i.e., that randomization worked.

Since we had no prior on the reach of the spillovers, we also explored larger radii (column 3). The coefficient on vouchers within 100-200m radius is statistically significant only in one round at the 5% and 10% in another, as shown in column 3. Again in the name of parsimony we decided not to include this variable in the IV regressions. The share of vouchers in the 200-300m radius to the specification in column 4 resulted non significant (not reported).

The effects of the individual discounts decrease by round 4. This is to be expected, since by this time roughly 80% of treatment group has connected to the grid, closer to the equilibrium connection rate. We exploit the fact that non-encouraged households started catching up with the encouraged group by round 4 of the survey to test whether these groups caught up also in outcomes. Showing that the effect of treatment in the non-encouraged group is similar to that on the encouraged group strengthens the policy implications derived from our study.

Table 3 reports our main experimental results. The dependent variable is $PM_{2.5}$ concentration (in logs) between 5pm and 7am[15]. The level of observation is the household-minute. To allow for arbitrary structure of the covariance matrix within household, we cluster the standard errors at the household level. By round 3, voucher recipients show drastic reductions compared to the non-encouraged group, with 67-73% lower $PM_{2.5}$ concentration[16]. When collapsing the data at the household level, the magnitude and significance remain unaltered. In rounds 4 and 5, the coefficients are closer to zero and not statistically significant, which we attribute to the control group catching up in electrification rate. Our sample is simply too small to pick up differences in $PM_{2.5}$ concentration with differences in electrification rates of around 10%. Figure 4 shows the reduced form results by hour of the day. The effects are larger from 5 to 10pm, decrease thereafter as most household members go to sleep around this time, and jump up again from 6 to 7am, when they wake up the next morning.

---

[15]In this sample we removed 17 households that reported average overnight $PM_{2.5}$ concentrations above 4 mg/day. Given the small overall number of households we take the conservative approach of excluding them. When they are included in the sample, the estimated reductions are much larger, but still within the confidence intervals in Table 4 (of the order of 90%).

[16]This figure is obtained from the reduced form coefficients: $e^{-1.119} - 1 \approx -0.67$; $e^{-1.316} - 1 \approx -0.73$.

Table 3B presents the IV results. The point estimate on connection is negative and large (implying a 95% reduction in overnight $PM_{2.5}$ concentration), but given the weak first stage, the standard errors are too large to draw any useful inference. The weakening of the first stage with time is consistent with the standard errors of the reductions blowing up in the later rounds, from .10 in the third round to 0.30 in the fourth and 1.8 in the fifth.

We present the CIC results to examine whether the non-encouraged group caught-up with the encouraged group in terms of $PM_{2.5}$ concentration by round 4 (given that the electrification rates were similar by then) and to explore the distribution of treatment effects.

The CIC estimator for the average treatment on the treated is -0.70. Given random group assignment, this is also the average treatment effect. The CIC estimator is consistent with the effects found among voucher recipients by round 3 (-0.67 to -0.73). This strengthens the internal validity of our findings to the extent that eliminating the differences in electrification rates led to eliminating the differences in overnight $PM_{2.5}$ concentration.

Figure 5 analyzes the variation in magnitude of treatment effects along the distribution of overnight $PM_{2.5}$ concentration. The percentage reductions in the outcome variable are of considerable size all throughout the distribution. However, there seems to be variability in the treatment effect along the distribution. The reduction is significant starting roughly from the 20th percentile, and the size of the effect starts increasing (becoming more negative) starting at the 60th percentile. This is consistent with the intuition behind our study setting: treatment effects are significant above a certain threshold of indoor air pollution, and higher polluters experience larger reductions.

4.2   Non-Experimental Estimates

In this section we show the results of non-experimental estimates to provide additional support to the experimental estimates from the previous section. We do not claim these estimates as causal, but we do show that there exists a solid relationship between electrification, kerosene consumption and overnight $PM_{2.5}$ concentration in a longitudinal setting. In this section we show that: (i) households that connect to the grid also reduce kerosene use and indoor air pollution; (ii) these change are not observed before households connect to the grid; (iii) these reductions are similar between groups irrespective of the timing of electrification; (iv) there is no reversion to pre-electrification levels in $PM_{2.5}$ concentration.

The non-experimental subsample shows a negative correlation between electrification and $PM_{2.5}$. In 2010, the (geometric) mean $PM_{2.5}$ concentration was 142 [95% Confidence Interval (CI) 122-165] $\mu g/m^3$ (N=201). In 2011, $PM_{2.5}$ concentration was 185 (149-230) $\mu g/m^3$ among non-connected households (N=119) and 128 (86-190) $\mu g/m^3$ among connected house-

holds (N=46). In 2012, $PM_{2.5}$ concentration was 108 (86-136) $\mu g/m^3$ among non-connected households (N=78) and 97 (75-125) $\mu g/m^3$ among connected households (N=90).[17]

In this subsection we will use electrification status at a given point in time to define groups, and we will use statistical tools to compare outcomes between those groups. None of the households in this sample had a connection by round 2 of the *EHEIPCER* survey (2010). T1 is the group of households that connected between rounds 2 and 3 of the survey (i.e., between 2010 and 2011). T2 is the group of houseoholds that connected between rounds 3 and 4 (between 2011 and 2012). T3 is the group of households that remain unconnected by 2012. Just a few households in the indoor air pollution sample connected to the grid after 2012. For the sake of readability of figures 6 and 7, we leave them in the control group.

Figure 6 shows the kernel density of $PM_{2.5}$ concentration between 1700 and 0700 hours by group. Panel (c) shows the kernel density estimates for T3 households. The density in 2010 and 2011 clearly overlap with each other. The two-sample Kolmogorov-Smirnov test for equality of distributions (KS) generates a p-value of .176; thus the null hypothesis of equality of distributions cannot be rejected at conventional confidence levels. The density in 2012 shows some differences with respect to the density in 2010; with an associated KS p-value< .001 in this case, thus rejecting the null hypothesis of equality of distributions. Panel (b) shows the same for T2 households. The densities overlap between 2010 and 2011 (KS p-value=.380), but the $PM_{2.5}$ density estimate shifted to the left 2012 (KS p-value< .001). For T1, the densities corresponding to the 2011 and 2012 measurements fall to the left of the 2010 distribution (KS p-value=.006 and < .001, respectively).

To allow for direct comparison with respect to the each group's baseline values, the variables in Figure 7 have been standardized by subtracting the baseline mean and dividing by the baseline standard deviation of their respective group. Panel (a) shows the change in average monthly expenditure on kerosene in 2011 compared to the 2010 levels by treatment arm, with 95% confidence intervals. There is no change in mean between the 2009 and the 2010 measurements for any of the groups. T3 does not show any change in mean kerosene expenditures in any of the surveys compared to 2010. T1 shows a large reduction between 2010 and 2011, which is maintained by 2012. T2 shows no change by 2011 (when the group is still off-grid), but a large reduction by 2012 (when the group connects to the grid).

Panel (b) shows the association with fine particulate matter. $PM_{2.5}$ concentration did not change significantly in the control group between baseline and follow-up. Households in T1

---

[17]Kerosene expenditure also shows a negative correlation with electrification in this subsample. In 2010, mean kerosene expenditure was 5.07 (4.57-5.61) US\$/month in 2010. In 2011, the figures were 5.54 (4.85-6.32) US\$/month for non-connected households and 3.87 (.15-99.42, due to only 2 non-zero observations) US\$/month. In 2012, mean kerosene expenditure was 5.65 (4.62-6.92) US\$/month among non-connected households and 2.52 (.93-6.81) US\$/month among connected households.

show significant reduction in average $PM_{2.5}$ concentration at the first and second follow-ups. Households in T2 show no reduction in $PM_{2.5}$ concentration by the first follow-up, but a significant reduction in mean $PM_{2.5}$ concentration by the second follow-up. This reduction is not statistically different from the average reduction experienced by households in T1.

This graph shows three striking facts. First, Both $PM_{2.5}$ and kerosene expenditure change when electrification status changes. Conversely, none changes if electrification status doesn't change (except for an *increase* in $PM_{2.5}$ among the Control group in 2011). Second, the average changes among households in T2 are similar to those experienced by households in T1. Third, the new (lower) levels of kerosene consumption and $PM_{2.5}$ observed for T1 in 2011 are maintained in 2012.

The results are presented in Table 4. Connection to the grid is associated with a 25-33% reduction in $PM_{2.5}$ concentration between 5pm and 7am. This estimate is consistent across specifications and strongly significant. Column (1) is a regression of $PM_{2.5}$ on connection and year fixed effects. Adding household fixed effects (column 2) does not alter the point estimate, but the standard errors blow up, suggesting that in this specification household fixed effects absorb too much of the variation in connection. In columns (3) and (4) we include subdistrict fixed effects and baseline characteristics in lieu of household fixed effects. The resulting point estimates implie a reduction of 22-26% in $PM_{2.5}$ concentration over the course of the study. In column (5) we test for differential treatment effects by round, and find that electrification led to substantial reductions in $PM_{2.5}$ concentration in rounds 3 and 4, but not in the last one. When the sample is restricted to the first four rounds of data, the coefficient on connection in a fixed effects estimation results -0.49, implying a 39% reduction in $PM_{2.5}$. This suggests something potentially different between households that connected in round 5 and the earlier rounds. In column (5) we replace connection status at round5 with connection status at round 4 and estimate the regression adding household fixed effects once again. The resulting coefficient is -0.33, significant at the 5% of confidence.

Using the data for T2 and T3 to test for differential pre-treatment trends in $PM_{2.5}$ and the 2010 *EHEIPCER* wave to test for differential pre-treatment trends in expenditure in kerosene or candles, use of wood, candles, we cannot reject the null hypothesis of parallel pre-treatment trends in any of the tests we performed.

4.3   Mechanisms

This subsection analyzes changes in traditional fuel use induced by electrification and suggests kerosene as the main channel through which electrification affected overnight $PM_{2.5}$ concentration. We report the effects of electrification on energy use in Table 6. Our find-

17

ings conform with the stylized fact that newly electrified households use electricity first and foremost for illumination.

The dependent variables in Table 6 indicate whether households use the particular energy source or not. Column 1 shows that voucher recipients are less likely to use kerosene and candles, although consistent with our previous discussion the coefficients are statistically significant only in round 2. However, the IV coefficients result negative and significant, at -0.33 for kerosene (p<.01) and -.10 for candles (p<.05). There are no significant changes in the use of car batteries, propane or wood. These sources are less important in the household's energy budget than kerosene, so detecting an effect would require larger sample sizes. As mentioned in the preceding section, similar patterns arise in the non-experimental sample.

4.4 Discussion on Effect Size

Taken at face value, these effects may seem too large. Given the first-stage and reduced-form coefficients, the implied IV estimator of the effect of electrification on overnight $PM_{2.5}$ concentration is given by:

$$\beta_{IV} = \frac{\beta_{RF}}{\beta_{FS}} \approx \frac{-0.60}{0.20} = -3.0 \tag{8}$$

, which implies a reduction in overnight $PM_{2.5}$ concentration of 95%[18]. However, in this section we argue that an effect this size is not out of the question and present some evidence in support of that claim. Note that the outcome of interest is $PM_{2.5}$ concentration between 5pm and 7am in the room of main use at night, in most cases the living room. In this room, and during this period, kerosene lamps arguably account for the largest share of $PM_{2.5}$ emissions. Note that Table 6 shows large decreases in the intensive and extensive margin of kerosene use, which should reflect on large drops in overnight $PM_{2.5}$ concentration.[19]

Next, we show that the drops implied by the model are consistent with the raw percentage changes in $PM_{2.5}$ concentration among households in our sample. We have successful $PM_{2.5}$

---

[18]In this case it is not possible to rely on the approximation $\frac{y_T - y_C}{y_C} \approx \frac{\partial ln y}{\partial connected} = \beta_1$, since $\beta_1 = -3$ is of considerable magnitude. The exact percentage change is given by:

$$\ln y_T - \ln y_C = \ln\left(\frac{y_T}{y_C}\right) = \beta_1$$
$$\frac{y_T}{y_C} = e^{\beta_1}$$
$$\frac{y_T - y_C}{y_C} = e^{-3} - 1 = -0.95$$

[19]Some emissions from cooking may filter during the day and linger in the living room, and, as we have seen, cooking practices did not change with electrification. However, there is no reason to believe that $PM_{2.5}$ concentration in the living room would depend more on filtrations from biomass during the day than from direct use of kerosene lamps in that room during the evenings.

measurements for 85 households in rounds 3 and 4. We calculated the change in $PM_{2.5}$ concentration for these households. Figure 8 shows the histogram of these changes.[20] Forty households show reductions of 50% or higher, 14 show reductions of up to 50%, and 30 households show increases in overnight $PM_{2.5}$ concentration. A few households show large percentage *increases*, but these large increases correspond to households with specially low levels of overnight $PM_{2.5}$ concentration at round 3. In levels, these increases are rather small. Conditional on showing an increase in $PM_{2.5}$, the average increase in levels was 0.051 mg/m^3, (95% CI 0.004 - 0.099 mg/m^3).

5 Implications for Health Outcomes

This section analyzes the health implications of the observed reductions in $PM_{2.5}$ concentration. First, we show that acute respiratory infections (ARIs) among children under the age of 6 were lower among voucher recipients. Next, we combine $PM_{2.5}$ concentration with time allocation data to construct measure of exposure to $PM_{2.5}$ for four typical household members (adult male, adult female, male child, female child), which allows gauging further health implications of the observed changes in $PM_{2.5}$ concentration.

5.1 Acute Respiratory Infections Among Children

Lower respiratory infections cause 2.8 million deaths globally in 2010 (Lozano et al., 2013) and thus they constitute a major public health concern. In this subsection we show that the reductions in overnight $PM_{2.5}$ concentration generated by household electrification had sizable effects on respiratory infections among children under six years old. The experimental sample includes 192 children in this age-range. Despite this relatively small sample size, there are large and statistically significant (at the 90% of confidence) reductions in the incidence of acute respiratory infections among children.

The dependent variable indicates whether the child had an episode of acute respiratory infection in the four weeks prior to the survey (self reported). When the explanatory variables are voucher, round, and their interactions, we find that vouchers led to a reduction of 16-18 percentage points at round 3 (significant at the 90%), depending on whether the regression controls for baseline characteristics and subdistrict fixed effects. The IV estimate on connection is -0.65, meaning that electrification reduced ARI incidence by 48 percentage points. However, it is important to notice that this result is significant at the 90% of confidence, and thus we can't reject much more modest reductions.

---

[20]Three households show percentage increases higher than 2.5. They are not included in the figure.

The point estimates for rounds 4 and 5 are not significant, consistent with the catching up argument. It is worth noting that it is not the case that ARIs bounced back up to their original levels, since ARI incidence reduced from 44% to 10% between rounds 3 and 4. So, consistent with the analysis of $PM_{2.5}$ concentration, this shows that the effects of electrification are similar irrespective of whether a household received a voucher.

5.2 $PM_{2.5}$ Exposure and Health Risks

In this subsection we construct measures of $PM_{2.5}$ exposure based on $PM_{2.5}$ concentration and time allocation data from the household survey. The resulting exposure measures are lower for the encouraged group (voucher recipients), but the gains are unequally distributed among household members. Adult males experience reductions in exposure to $PM_{2.5}$ nearly twice as large as adult females, mostly due to inequality in time spent cooking, where $PM_{2.5}$ concentration is highest, and outdoors, where $PM-2.5$ concentration lowest. It is important to keep in mind that given imperfect compliance with voucher assignment, the reduced form coefficients (the effect of voucher on overnight $PM_{2.5}$ concentration) constitute lower bounds of the effects of electrification.

We start by defining exposure to a pollutant as the amount (in mg) of pollutant that effectively makes its way into a person's respiratory system. (As discussed earlier, $PM_{2.5}$ enters the deep lung and the bloodstream.) For a particular activity conducted for a given amount of time, exposure is estimated as the product of the concentration in the environment where this activity took place multiplied by the inhalation rate while performing said activity. Daily exposure can be estimated by adding over all activities performed during the day, as in the following equation:

$$\text{Exposure} = \sum_{j=1}^{J} \text{time}_j \times \text{concentration}_j \times \text{inhalation}_j \tag{9}$$

where time_j is time spent in activity j (hours), concentration_j is $PM_{2.5}$ concentration in the room while performing activity j (mg/m^3), and inhalation_j is the estimated inhalation rate while performing activity j ($m^3/hour$).

The detailed information on time allocation of the household members gives our study an important advantage over the typical study in this literature.[21] It allows to estimate fairly accurately the time individuals spent indoors and, moreover, the type of activity, they were performing. Knowing the type of activity and the time allocated to said activity allows

---

[21]This is even more so when compared to outdoor pollution studies, in which time spent outdoors is an unobservable variable

inputting an average inhalation rate adequate for that particular activity during that time, instead of just inputting the average intake per day. For a given pollutant concentration, this leads to better estimates of the amount of pollutants that effectively makes its way into a persons respiratory system.

The time allocation data was collected for up to four household members: the household head, his or her spouse, and up to two school-age children. This allows estimating $PM_{2.5}$ exposure for four "synthetic individuals": adult female, adult male, female child, and male child. Table 8 presents average time allocation in four type of activities for each of our synthetic individuals. The male and female heads report 8.9 hours of sleep per day, while the children report 9.5-10 hours of sleep per day. Time at home during the evening (awake) is similar for all members (slightly lower for the children, who sleep 0.5-1.0 hours more than the household heads). The starkest differences are observed in time spent in the kitchen and time spent outside the home. While female head reports 2.5 hours per day in the kitchen, the male head reports spending an average of just five minutes. On the other hand, the female head reports spending an average of 2.7 hours outside the home, while the male head reports 7.8 hours (consistent with the length of a work day).

The differences in time allocation that arise from this analysis already suggest that adult females are more exposed to $PM_{2.5}$ (as well as other pollutants) since they spend considerably more time in the kitchen than any other household member. On the other hand, males spend almost one third of the time outside the home. The main activity is farming and walking to and from the farm, where it can be safely assumed that exposure to $PM_{2.5}$ is negligible.

Next, we make explicit the assumptions about the $PM_{2.5}$ concentration in the environments where these activities were likely conducted. As shown earlier, average $PM_{2.5}$ concentration in the living room during the evenings is 0.40 mg/m^3. We take this as representative of any room in the household, except the kitchen, between 1700 hours in the evening and 0700 the next morning. Based on the subsample of households for which we have 3-day measurements, we estimate the average $PM_{2.5}$ concentration in the living room during daytime (from 0700 to 1700 hrs) to be 0.26 mg/m^3. We take this as representative of the rooms in the household during daytime, again with the exception of the kitchen. Since we did not collect data on $PM_{2.5}$ concentration in the kitchen, we use 0.90 mg/m^3, which corresponds to average $PM_{2.5}$ in the kitchen in Guatemalan households (Northcross et al., 2010). This figure seems an adequate assumption in our context since it corresponds to a neighboring region where households also rely of fuelwood for cooking. This makes our exposure estimates adequate for households that rely on wood for cooking, and even conservative given that its not uncommon to find cases where the average concentration is above 2.0 mg/m^3. We assume household members will not be exposed to $PM_{2.5}$ whenever they are not home.

This assumption seems not to be too restrictive for the population in our study setting, since most of the time outside the home is spent in outdoors activities like farming, and very little time is spent conducting activities outside the home that suggest exposure to $PM_{2.5}$ (e.g. visiting friends at night).

The third and final component in (9) is the inhalation rate. Since inhalation rate depends on age, we estimated it for the sample averages: 43 for the female head, 47 for the male head, 11 for the female child, and 13 for the male child. Air inhalation rates per activity are based on the EPA Exposures Handbook (EPA, 2011). Most activities conducted at home can be classified as "light activity tasks" by the EPA. Light activities include cooking, washing dishes, ironing, watching TV, doing desk work, writing and typing, and walking at a speed of up to 2.5 mph (2.9 km/h). The average inhalation rate for these activities is 0.78 $m^3/hour$, while the average air inhalation rate while sleeping is 0.30 $m^3/hour$, again simliar for the four synthetic individuals. The inhalation rate for activities conducted outside the home will vary greatly, depending on the intensity of these activities. For instance, walking to work could be classified as light or medium intensity, depending on the speed at which the person is walking. Farming, on the other hand, could be classified as medium to high intensity, but if lunch breaks would be light activity. However, the assumption made earlier about $PM_{2.5}$ concentration being zero outside the home makes the inhalation rates of these activities irrelevant for total exposure.

With these three components we estimated exposure rates for the four synthetic individuals. Estimated exposure measures are highest for the female head, at 5.68 mg/day, and lowest for the male head, at 3.20 mg/day. The exposure measures for children lie in between, with females 4.23 mg of $PM_{2.5}$ per day males to 3.72 mg/day. Taken plainly as units of $PM_{2.5}$, these concentrations are equivalent to 8.0 cigarettes a month for the male head, 14.2 for the female head, 10.6 for the female child, and 9.3 for the male child[22]. The scientific evidence is yet inconclusive as to whether generated by $PM_{2.5}$ cigarette is worse or not than that generated by kerosene combustion.

The changes in exposure are large for all members (all of them above 30%), but these gains are unequally distributed across household members. The male head benefits the most, with a reduction in exposure of 59%, while the female head benefits the least, with a reduction of 33%. As pointed above, these differences owe to females spending more time than males in the kitchen, where pollutant concentration is highest, while males spend more time than females outside the home, where pollutant concentration is lowest.

To date, there are no dose-response functions linking exposure to $PM_{2.5}$ from kerosene combustion to health outcomes. However, Pope III et al. (2011) presents an estimate of a

---

[22]One cigarette is estimated to have 12mg of $PM_{2.5}$.

dose-response function linking $PM_{2.5}$ from first- and second-hand tobacco smoking to lung cancer and cardiovascular diseases. In Panel (b) of Table 8 we present the relative risks that would be associated to the exposure levels found in Panel (a) if the health effects of $PM_{2.5}$ from kerosene combustion were similar to those from tobacco smoking. It is worth noting that the dose-response function estimated by Pope III et al. (2011) is non-linear and it has support in exposures of from 0.18 to 0.90 mg/day and then above 18 mg/day (but not between 0.90 and 18mg/day, while our estimates range from 3.2 to 5.7) so we need to rely on the linear interpolation of the values up to 18mg/day. This caveat does not seem a strong weakness since the linearization is highly accurate in this neighborhood, with R-squared values of 0.99 (lung cancer), 0.96 (ischemic heart disease), 0.86 (cardiovascular disease), and 0.80 (cardiopulmonary disease).

The changes in exposure are associated with a decrease in relative risk of lung cancer (compared to a person with no exposure to $PM_{2.5}$) from 4.0 to 3.1 for the female head, nearly a 25% reduction. The relative risk for the male head falls by 33%, while the reduction for female child is 25% and for the male child is almost 30%. The estimated reductions in the relative risk of ischemic heart disease, cardiovascular disease, and cardiopulmonary disease are between 3 to 4%. Consistent with the results in lung cancer, these changes are higher for adult males, but the differences at these levels of exposure are relatively small.

6   Conclusions

This paper provides the first experimental evidence that electricity leads to important improvements in welfare, through substantial, immediate and sustained improvements in indoor air quality. To this end, we pair a clean experimental design with an exceptionally rich dataset with minute-by-minute data on fine particulate matter ($PM_{2.5}$) concentration and time allocation for a sample of households in Northern El Salvador during a recent electrification program. In this program, households had to pay a $100 fee to connect to the grid (20% of the sample annual per capita income). By randomly allocating discount vouchers towards this fee we generated exogenous variation in the cost of connection to the grid.

Given imperfect compliance with voucher assignment, our experiment produced a lower bound for reductions in average $PM_{2.5}$ concentration, of 67%, with the IV estimates suggesting reductions of the order of 95%. The most salient mechanism behind the improvements in IAP is a substitution away from kerosene lighting, while there are no discernible changes in the use of other traditional lighting sources or in cooking practices. Given that the mechanism is clear in this context, our results suggest that other clean artificial lighting technologies, like solar lamps, could have similarly strong effects on indoor air pollution in

households that are too isolated for grid electrification to be feasible.

As a result of this drop overnight $PM_{2.5}$ concentration, we see large and significant falls in acute respiratory infections (ARI) among children under the age of 6. Our IV estimates suggest that connected households have 65% lower incidence of ARI. To gauge further implications on health for population over the age of 6, the observed changes in $PM_{2.5}$ concentration, together with time-use data, and complementary $PM_{2.5}$ data from Northcross et al. (2010) were used to estimate the change in daily $PM_{2.5}$ exposure. We find that reductions in average exposure to $PM_{2.5}$ are large for all members, but they are distributed unequally among household members. Adult males typically benefit the most from the reductions in $PM_{2.5}$ brought about by electrification, with 59% lower exposure. Adult females benefit the least, with reductions of 33%. The figures for children are 46% (males) and 39% (females). To assess the magnitude of the health effects, we input these figures into the dose-response function estimated by Pope III et al. (2011). The implied risk-ratio for lung cancer fall dramatically, from 4.0 to 3.1 for adult females, from 2.7 to 1.8 for adult males, with the respective figures for children falling from 3.2 to 2.4 (females) and from 3.0 to 2.1 (males). The main caveat in this analysis is that these risk ratios rely on the assumption that the health effects of $PM_{2.5}$ from kerosene combustion are similar to those of cigarette smoking. The exact figures depend on the health effects of $PM_{2.5}$ from kerosene combustion relative to those from cigarette smoking. However, given the magnitudes of exposure and the reductions we find, we argue that the health effects will be large.

This paper also contributes to the environmental health literature by providing evidence of a strong, positive relationship between kerosene use and $PM_{2.5}$ concentration in a setting with high reliability on biomass for cooking (70% of the sample households). This is important given the still scarce evidence on the relationship in the field between kerosene use and IAP measures. Despite the large improvements in indoor air quality brought about by electrification, $PM_{2.5}$ concentration in these households is still high, due to the use of fuelwood for cooking. In fact, this is the reason behind the higher exposure levels among females than males. This highlights the importance of advancing our understanding on the adoption and use of improved cookstoves.

We also examined heterogeneity in treatment effect with CIC, a non-parametric approach that allows estimating the full distribution of treatment effects. Households up to the 20th percentile (lowest polluters) do not benefit from electrification[23], but there are large and significant reductions for all households from the 20th percentile onwards. Its also noticeable that the reduction starts getting larger among households in the 60th percentile and are

---

[23]These households have extremely low $PM_{2.5}$ concentrations at round 3, so any small increase at round 4 (even within statistical error) results in a large percentage increase.

largest for the 15% of households with the highest pre-electrification $PM_{2.5}$ concentration.

References

Athey, Susan and Guido W Imbens. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica* 74 (2):431–497.

Barnes, Douglas F. 2007. *The challenge of rural electrification: strategies for developing countries.* Earthscan.

Barron, Manuel and Maximo Torero. 2013a. "Household Electrification and Time Allocation." *mimeo* .

———. 2013b. "Kerosene Use and Indoor Air Pollution: Longitudinal Evidence from Northern El Salvador." *mimeo* .

Bensch, Gunther, Jochen Kluve, and Jörg Peters. 2011. "Impacts of rural electrification in Rwanda." *Journal of Development Effectiveness* 3 (4):567–588.

Bernard, Tanguy. 2012. "Impact analysis of rural electrification projects in sub-Saharan Africa." *The World Bank Research Observer* 27 (1):33–51.

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

Chay, Kenneth Y and Michael Greenstone. 2003. "The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession." *The quarterly journal of economics* 118 (3):1121–1167.

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

Chowdhury, Zohir, Rufus D Edwards, Michael Johnson, Kyra Naumoff Shields, Tracy Allen, Eduardo Canuz, and Kirk R Smith. 2007. "An inexpensive light-scattering particle monitor: field validation." *Journal of Environmental Monitoring* 9 (10):1099–1106.

Dinkelman, Taryn. 2011. "The effects of rural electrification on employment: New evidence from South Africa." *The American Economic Review* 101 (7):3078–3108.

Edwards, Rufus, Kirk R Smith, Brent Kirby, Tracy Allen, Charles D Litton, and Susanne Hering. 2006. "An inexpensive dual-chamber particle monitor: laboratory characterization." *Journal of the Air & Waste Management Association* 56 (6):789–799.

EPA. 2011. "Exposure Factors Handbook: 2011 Edition." Tech. rep., US Environmenal Protectin Agency - National Center for Environmental Assessment, Washington, DC.

Fan, Cheng-Wei and Junfeng Jim Zhang. 2001. "Characterization of emissions from portable household combustion devices: particle size distributions, emission rates and factors, and potential exposures." *Atmospheric environment* 35 (7):1281–1290.

Grogan, Louise and Asha Sadanand. 2012. "Rural Electrification and Employment in Poor Countries: Evidence from Nicaragua." *World Development* .

Hanna, Rema, Esther Duflo, and Michael Greenstone. 2012. "Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves." Tech. rep., National Bureau of Economic Research.

Hanna, Rema and Paulina Oliva. 2011. "The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City." Tech. rep., National Bureau of Economic Research.

Independent Evaluation Group. 2008. "The Welfare Impacts of Rural Electrification: A Reassessment of the Costs and Benefits." Tech. rep., World Bank.

International Energy Agency. 2011. "World Energy Outlook 2011."

Khandker, Shahidur R, Douglas F Barnes, and Hussain A Samad. ???? "Forthcoming.– Welfare Impacts of Rural Electrification: A Panel Data Analysis from Vietnam." *Economic Development and Cultural Change* .

Lam, Nicholas L, Yanju Chen, Cheryl Weyant, Chandra Venkataraman, Pankaj Sadavarte, Michael A Johnson, Kirk R Smith, Benjamin T Brem, Joseph Arineitwe, Justin E Ellis et al. 2012a. "Household light makes global heat: high black carbon emissions from kerosene wick lamps." *Environmental science & technology* 46 (24):13531–13538.

Lam, Nicholas L, Kirk R Smith, Alison Gauthier, and Michael N Bates. 2012b. "Kerosene: a review of household uses and their hazards in low-and middle-income countries." *Journal of Toxicology and Environmental Health, Part B* 15 (6):396–432.

Lim, Stephen S, Theo Vos, Abraham D Flaxman, Goodarz Danaei, Kenji Shibuya, Heather Adair-Rohani, Markus Amann, H Ross Anderson, Kathryn G Andrews, Martin Aryee et al. 2013. "A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010." *The lancet* 380 (9859):2224–2260.

Litton, Charles D, Kirk R Smith, Rufus Edwards, and Tracy Allen. 2004. "Combined optical and ionization measurement techniques for inexpensive characterization of micrometer and submicrometer aerosols." *Aerosol Science and Technology* 38 (11):1054–1062.

Lozano, Rafael, Mohsen Naghavi, Kyle Foreman, Stephen Lim, Kenji Shibuya, Victor Aboyans, Jerry Abraham, Timothy Adair, Rakesh Aggarwal, Stephanie Y Ahn et al. 2013. "Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010." *The Lancet* 380 (9859):2095–2128.

Masera, Omar R, Rodolfo Díaz, and Víctor Berrueta. 2005. "From cookstoves to cooking systems: the integrated program on sustainable household energy use in Mexico." *Energy for Sustainable Development* 9 (1):25–36.

Mobarak, Ahmed Mushfiq, Puneet Dwivedi, Robert Bailis, Lynn Hildemann, and Grant Miller. 2012. "Low demand for nontraditional cookstove technologies." *Proceedings of the National Academy of Sciences* 109 (27):10815–10820.

Northcross, Amanda, Zohir Chowdhury, John McCracken, Eduardo Canuz, and Kirk R Smith. 2010. "Estimating personal PM2. 5 exposures using CO measurements in Guatemalan households cooking with wood fuel." *Journal of Environmental Monitoring* 12 (4):873–878.

Pokharel, Shaligram. 2003. "Promotional issues on alternative energy technologies in Nepal." *Energy Policy* 31 (4):307–318.

Pope III, C Arden, Richard T Burnett, Michelle C Turner, Aaron Cohen, Daniel Krewski, Michael Jerrett, Susan M Gapstur, and Michael J Thun. 2011. "Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: shape of the exposure–response relationships." *Environmental health perspectives* 119 (11):1616.

Röllin, HB, A Mathee, N Bruce, J Levin, and YER Von Schirnding. 2004. "Comparison of indoor air quality in electrified and un-electrified dwellings in rural South African villages." *Indoor Air* 14 (3):208–216.

Schare, Stuart and Kirk R Smith. 1995. "Particulate emission rates of simple kerosene lamps." *Energy for Sustainable Development* 2 (2):32–35.

Smith, Kirk R, Gu Shuhua, Huang Kun, and Qiu Daxiong. 1993. "One hundred million improved cookstoves in China: How was it done?" *World Development* 21 (6):941–961.

Troncoso, Karin, Alicia Castillo, Omar Masera, and Leticia Merino. 2007. "Social perceptions about a technological innovation for fuelwood cooking: Case study in rural Mexico." *Energy Policy* 35 (5):2799–2810.

van de Walle, Dominique, Martin Ravallion, Vibhuti Mendiratta, and Gayatri Koolwal. 2013. "Long-term impacts of household electrification in rural India." *World Bank Policy Research Working Paper* (6527).

Wallmo, Kristy, Susan K Jacobson et al. 1998. "A social and environmental evaluation of fuel-efficient cook-stoves and conservation in Uganda." *Environmental Conservation* 25 (2):99–108.

WHO. 2006. "Fuel for life: household energy and health." *World Health Organization, Geneva* .

World Bank. 2007. "Clean Energy for Development Investment Framework: The World Bank Group Action Plan." Tech. rep., The World Bank.

———. 2009. Tech. rep., The World bank.

7   Appendix: Changes in Changes

As discussed earlier, treatment effect is likely to vary depending on each household's pre-electrification $PM_{2.5}$ concentration, which would violate the assumption of constant treatment effect. To examine possible violations to this assumption, we employ changes-in-changes (CIC), a non-linear, non-parametric generalization of differences-in-differences (DID) developed by Athey and Imbens (2006) that, among other features, does not impose the assumption of constant treatment effect. In this section we summarize the main points of CIC and towards the end we draw a parallel with this study.

The intuition behind this methodology is straightforward when compared to DID. The DID approach uses the change in means in the control group to estimate the counterfactual change in means in the treatment group and derive an average treatment effect on the treated group. Under less restrictive assumptions than DID, CIC uses the change in the full distribution of the outcome variable in the control group to estimate the counterfactual change in the distribution in the treatment group, and derives not only the average treatment effect on the treated, but also the distribution of those treatment effects. This feature makes it especially appealing for our purposes. Unlike DID, where time and group dimensions are treated symmetrically, CIC allows the effects of time and treatment to differ systematically across groups. It also estimates the entire counterfactual distribution of effects of the treatment on the treated and of the treatment on the control group.[24]

The model assumes that the outcome variable Y for an individual in the absence of the intervention is given by

$$Y^N = h(U, T) \tag{10}$$

, where h is a "production function" that depends on the period T, and an unobservable variable U. Y^N is the outcome of interest under no treatment. In this case, treatment assignment will be defined by voucher allocation. We can think of the unobservable U as being household's ability and preferences toward health or a clean environment.

The identification in CIC relies on the assumption that in each time period the production function is monotone in an unobservable U. The distribution of U may in fact differ between treatment and control groups but, absent treatment, households with the same U must have the same value of Y^N irrespective of their voucher assignment. Note that Y^N is the counterfactual value of Y_{11} under no treatment.

Formally, CIC is based on the following assumptions[25]:

Assumption 1 The model: *The outcome of an individual in the absence of interventions satisfies the relationship $Y^N = h(U, T)$*

Assumption 2 Strict monotonicity: *The production function $h(u, t)$, where $h : \mathbb{U} \times \{0, 1\} \to \mathbb{R}$, is strictly increasing in u for $t \in \{0, 1\}$*

Assumption 3 Time invariance within groups $U \perp T | G$

---

[24]In addition, CIC allows for the possibility that the treated group took-up treatment because of greater benefits than the control group. We do not make use of this property, since vouchers were randomly allocated.

[25]Athey and Imbens (2006) shows that DID can be obtained by adding assumptions...

28

Assumption 4 Support. $\mathbb{U}_1 \subseteq \mathbb{U}_0$.

Assumption 1 requires outcomes not to depend on the group indicator directly and that all relevant unobservables can be captured in a single index U. In this study, groups are defined by voucher allocation. Given random group assignment, there is no reason to believe that outcomes would depend directly on the group indicator (other than because vouchers increase the probability of electrification, and that will change the production function from h^N to some other function h^I). Assumption 2 requires higher unobservables to correspond to strictly higher outcomes. Households with higher ability or higher valuation for health could have better indoor environment, by using less kerosene or better ventilation rates, for instance. The assumption about the unobservables cannot be tested but it is consistent with basic intuition, and it allows for non-additive structures that arise in economic models. Assumption 3 requires the population of agents within each group not to change over time. In other words, it requires any differences between the groups to be stable, which seems feasible given random voucher allocation. This allows trend in one group to be used to estimate the counterfactual trend in the in the other. Note that under this assumption any change in the variance of outcomes over the time period will be attributed to changes over time in the production function.[26] Finally, assumption 4 implies that $\mathbb{Y}_{10} \subseteq \mathbb{Y}_{00}$ and $\mathbb{Y}_{11}^N \subseteq \mathbb{Y}_{01}$, i.e. we can only estimate the distribution function of Y_{11}^N on the support $\mathbb{Y}_{01}$. The minute-by-minute dataset on $PM_{2.5}$ concentration gives adequate support in both groups and in both time periods.

We start by defining $F_Y^{-1}(q)$. For $q \in [0, 1]$ and for a random variable Y with compact support $\mathbb{Y}$,

$$F_Y^{-1}(q) = \inf\{y \in \mathbb{Y} : F_Y(y) \geq q\} \tag{11}$$

Athey and Imbens (2006) show that under the assumptions discussed above the distribution of F_Y^{-1} is identified and that the counterfactual value density Y^N for the treated group in the second period is given by:

$$F_{Y^N,11}(y) = F_{Y,10}\left(F_{Y,00}^{-1}\left(F_{Y,01}(y)\right)\right) \tag{12}$$

The mechanics of this calculation are straightforward. Start by taking a value of the outcome variable y_0 in the treated group before treatment, corresponding to a percentile q_0 in $F_{Y,10}$. We want to find the counterfactual distribution of Y_{11} for the treated group, F_{11}^N, which we do not observe. To estimate F_{11}^N, first find the quartile q_0' corresponding to y_0 in the CDF of Y_{00}. Next, measure the change in the outcome variable between periods 0 and 1 for the percentile q_0' and note the change in y_0, Δy_0. Then calculate $y_1 = y_0 + \Delta y_0$ and attribute it to the original q_0. The point (y_1, q_0) lies on the counterfactual CDF of Y_{11}. The full counterfactual CDF F_{11}^N is obtained by repeating this procedure for all values of y in the domain.

---

[26]As noted by Athey and Imbens (2006), this contrasts with DID with full independence (which rules out changes in variance), and DID with mean independence (which ignores these changes).

The average treatment effect in CIC is then given by[27,28]:

$$\tau^{CIC} \equiv \mathbb{E}\left[Y_{11} - Y_{11}^N\right] \tag{13}$$

$$= \mathbb{E}\left[Y_{11}\right] - \mathbb{E}\left[F_{Y,10}\left(F_{Y,00}^{-1}\left(F_{Y,01}\left(y\right)\right)\right)\right] \tag{14}$$

In section 4.1 we show that the connection rate among the non-encouraged group (i.e. households that did not receive discount vouchers) started catching up with that of the encouraged group by round 4 of the survey. We exploit this fact to analyze whether overnight $PM_{2.5}$ concentration in the non-encouraged group also caught up with the encouraged group.

As discussed above, vouchers gave incentives to adopt electricity in rounds 2 and 3 instead of in later rounds. Thus, vouchers can be thought of as an incentive *not to connect* in round 4. Effectively, the electrification rate among voucher recipients barely changed between rounds 3 and 4. On the contrary, non recipients had a faster electrification rate between those rounds, almost catching up with voucher recipients. Note that this change in electrification rates was also exogenously induced by random voucher allocation. But in this case we should think of voucher recipients as "control" group between rounds 3 and 4, and non-recipients as "treatment" group in that period. Hence, we will use voucher recipients in rounds 3 and 4 to estimate $F_{Y,00}$ and $F_{Y,01}$ and non-recipients to estimate $F_{Y,10}$ and $F_{Y,11}$, respectively.

---

[27]This contrasts with the treatment effect in DID: $\tau^{DID} = \mathbb{E}\left[Y_{11}\right] - \left(\mathbb{E}\left[Y_{10}\right] - \left(\mathbb{E}[Y_{01}] - \mathbb{E}\left[Y_{00}\right]\right)\right)$

[28] Athey and Imbens (2006) draw a parallel with quantile estimation, which they name quantile differences in differences (QDID). The authors show that CIC has more desirable properties than QDID, so we will concentrate on CIC.

Figures And Tables

Figure 1. El Salvador: Socio-economic Status, Electrification, and
Use of Traditional Fuels, 2007

1A - District SES, El Salvador 2007

Quartile 1 (poorest)
Quartile 2
Quartile 3
Quartile 4 (wealthiest)
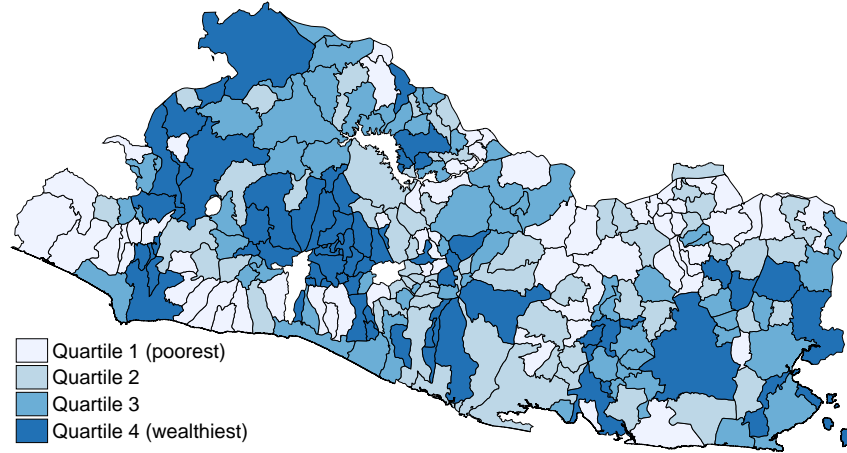
1B - Electrification rates (%) El Salvador 2007

[30.3,78]
(78,85.4]
(85.4,90.1]
(90.1,98.9]
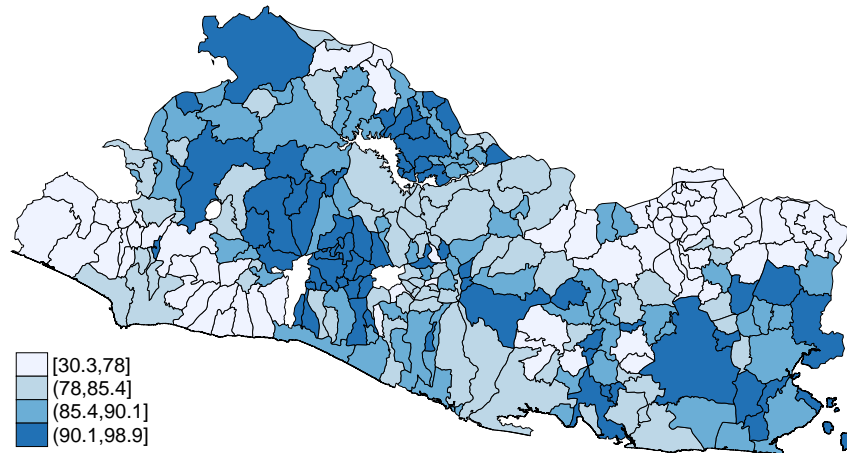
1C - Use of traditional fuels for cooking or ligthing (%) El Salvador 2007

[2.4,41]
(41,57.3]
(57.3,74.3]
(74.3,95.5]
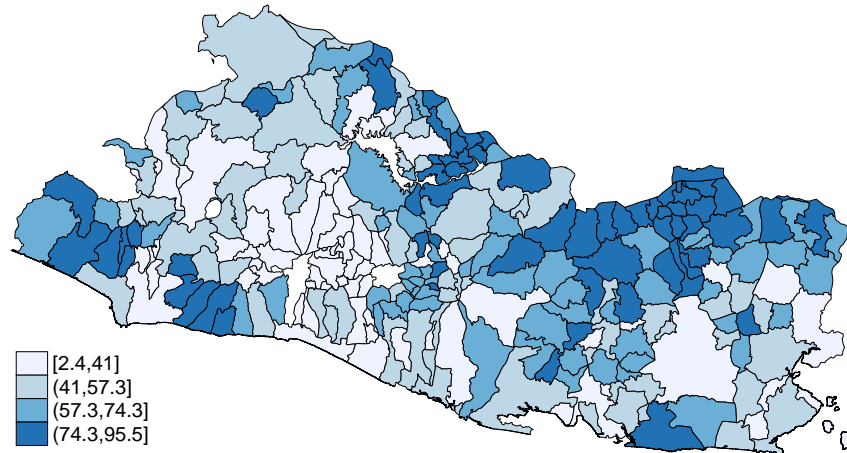
Source: 2007 National Census

Figure 2. Monthly expenditure in kerosene and PM$_{2.5}$ concentration
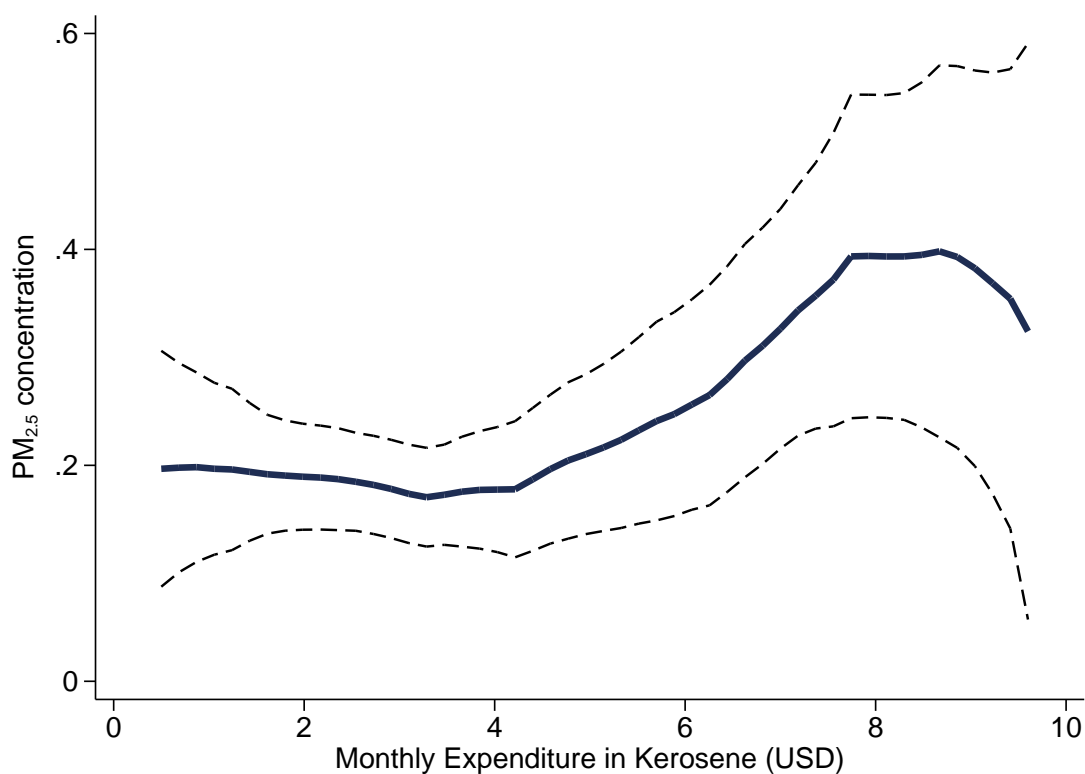
The figure shows a descriptive relationship between monthly expenditure in kerosene and fine particulate matter (PM$_{2.5}$) concentration ($\mu g/m^3$) from 1700-0700hrs. PM$_{2.5}$ concentration is representative of the room where most household members spent most of their time awake during the evening (typically the living room). For details on the PM$_{2.5}$ measurement protocol, see appendix.

Figure 3. Connection Rates by Treatment Arm
(experimental sample)
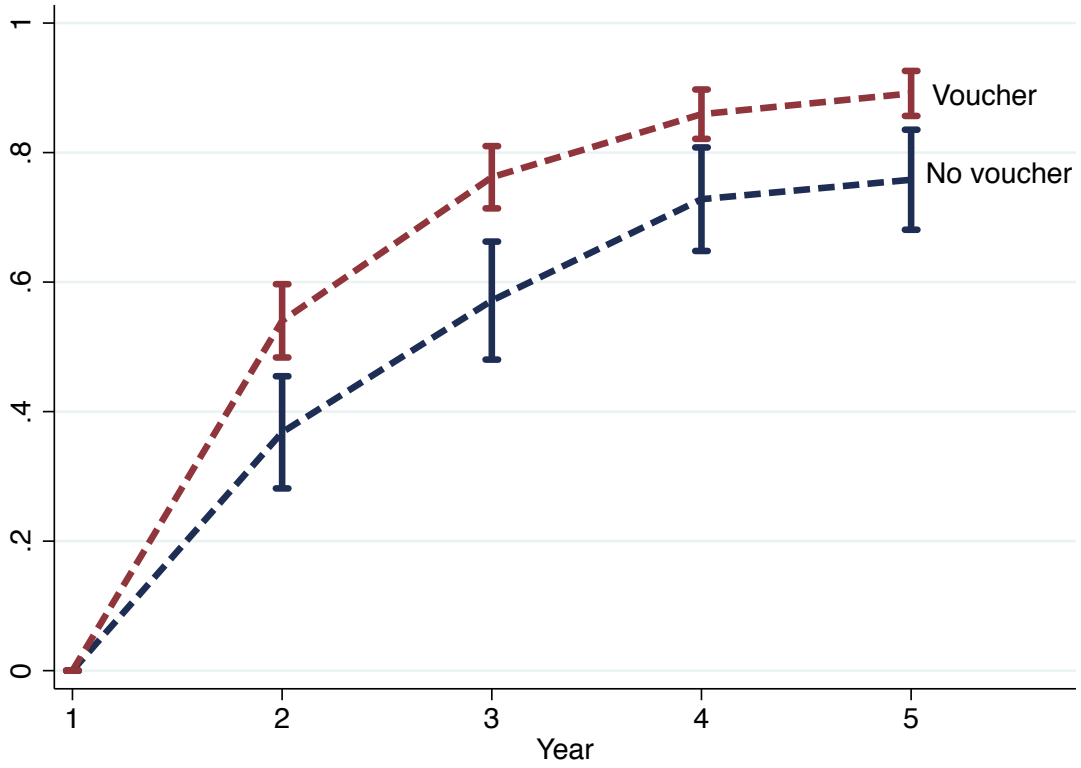
This figure shows the evolution of connection rates between voucher recipients (encouraged group) and non-recipients (non-encouraged) over the study period.

Figure 4.Voucher allocation and PM$_{2.5}$ concentration by hour of the day (95% Confidence Intervals)
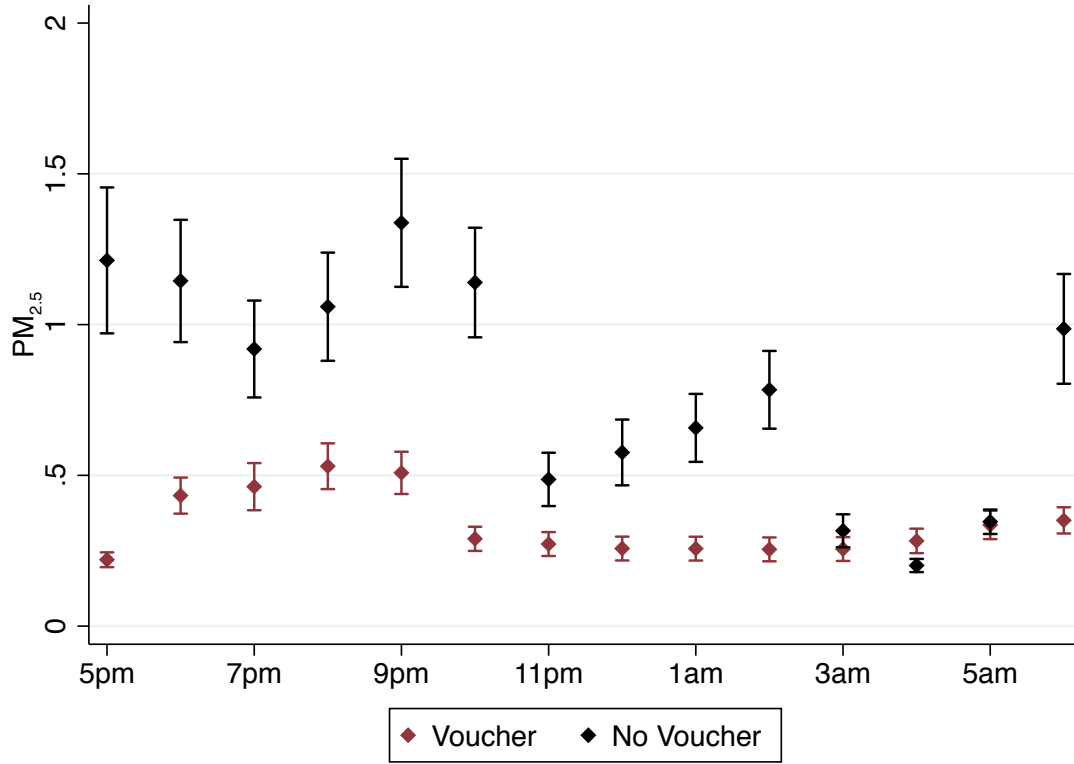
Figure 5. Heterogeneity in Treatment Effect, CIC Model
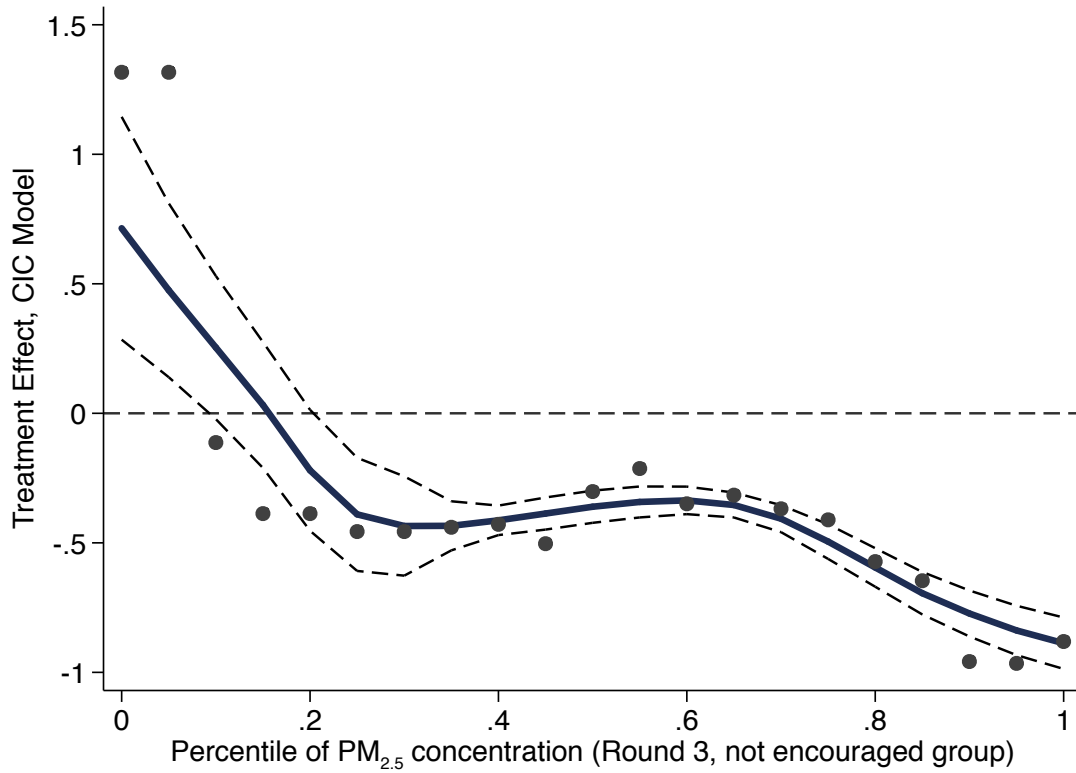
The figures shows the treatment effect estimated by the Changes-in-Changes model. The vertical axis measures the percentage change (1=100%) between rounds 3 and 4 implied by the Changes-in-Changes model. The horizontal axis indicates the percentile of overnight $PM_{2.5}$ concentration in round 3 for the non-encouraged group (i.e. households that did not receive vouchers at baseline). The solid line represents the local polynomial estimator of the effect, and the dotted lines correspond to its 95% confidence interval.

Figure 6. Kernel densities for overnight PM$_{2.5}$ concentration, by electrification status (non-experimental sample)

Electrification and PM$_{2.5}$ Concentration (μg/m^3, 1700-0700hrs)
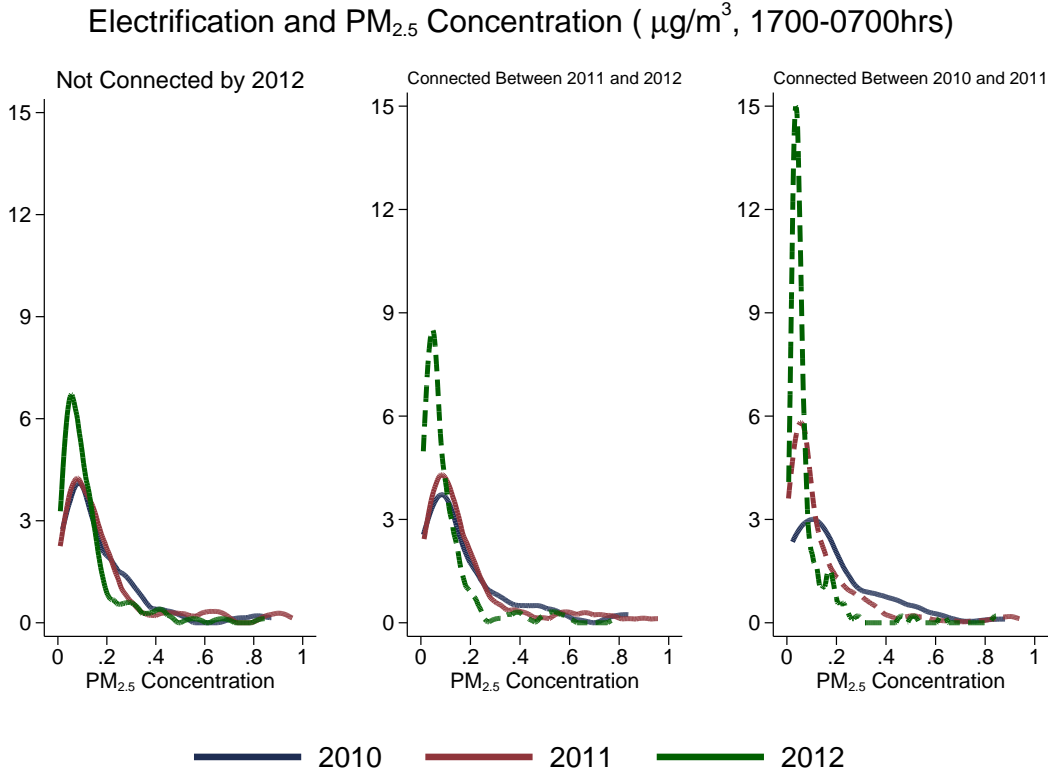
The figure shows the estimated density of overnight fine particulate matter (PM$_{2.5}$) concentration ($\mu g/m^3$) by electrification status of the household, from 2010 to 2012, on the non-experimental air quality subsample (see main text for details). The left panel shows PM$_{2.5}$ concentration for households that remained off-grid during the study period. The center panel shows PM$_{2.5}$ concentration for households that had connected only by 2012 (green line). The panel on the right shows the respective densities for households that connected between 2010 and 2011 (green and red lines). Solid lines indicate that the group was had not connected to the grid yet, dashed lines indicate that the group had connected to the grid.

Figure 7. Electrification, Kerosene use, and PM$_{2.5}$
(non-experimental sample)
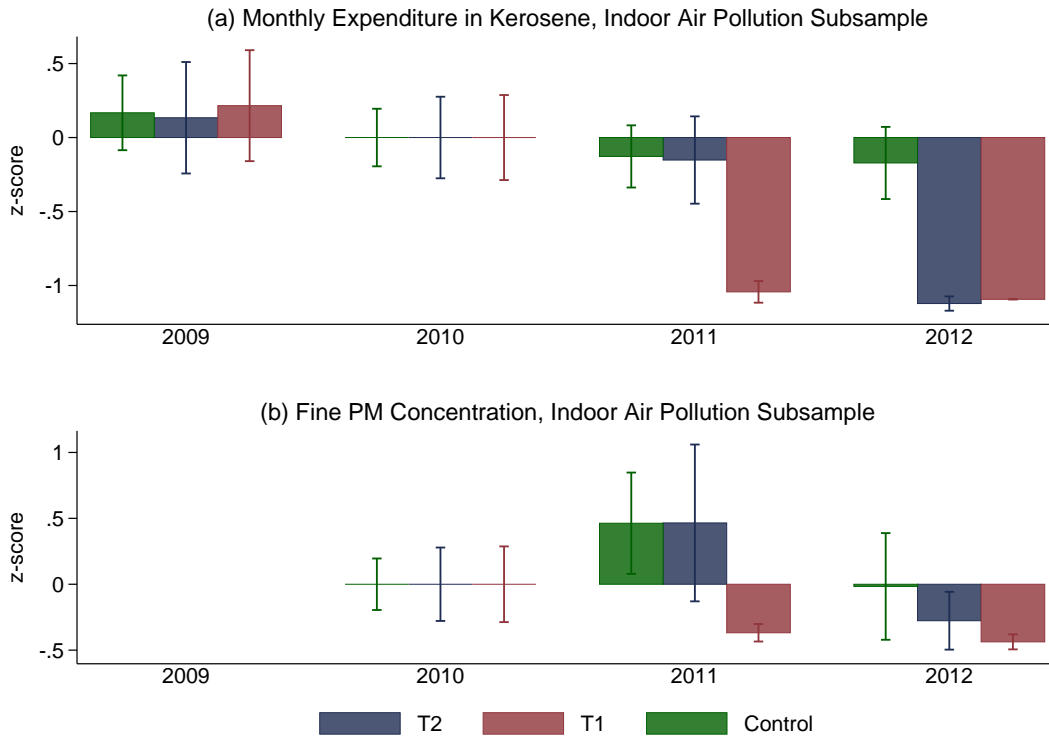
This figure shows the changes in kerosene expenditure and PM$_{2.5}$ in the non-experimental air quality subsample with respect to its 2010 values (see main text for details on the non-experimental sample). The green group ("Control") remained off-grid during the course of the study. No significant changes are observed for this group with respect to its 2010 values (the confidence interval includes zero in all but one case). The red group ("T1") is formed by households that got a grid connection between the 2010 and the 2011 surveys. Both its kerosene consumption and its overnight PM$_{2.5}$ concentration are significantly lower with respect to their baseline values. Comparing the red bars in 2011 to 2012 shows that these changes persist in time. The blue group ("T2") is formed by households that connected to the grid between the 2011 and the 2012 surveys. While the group was off-grid, its mean kerosene consumption and PM$_{2.5}$ concentration are not significantly different than the figures for the green group. While it is connecte,d its mean kerosene consumption are not significantly different than the red group.

Figure 8. Unadjusted reductions in overnight $PM_{2.5}$ concentration between rounds 3 and 4
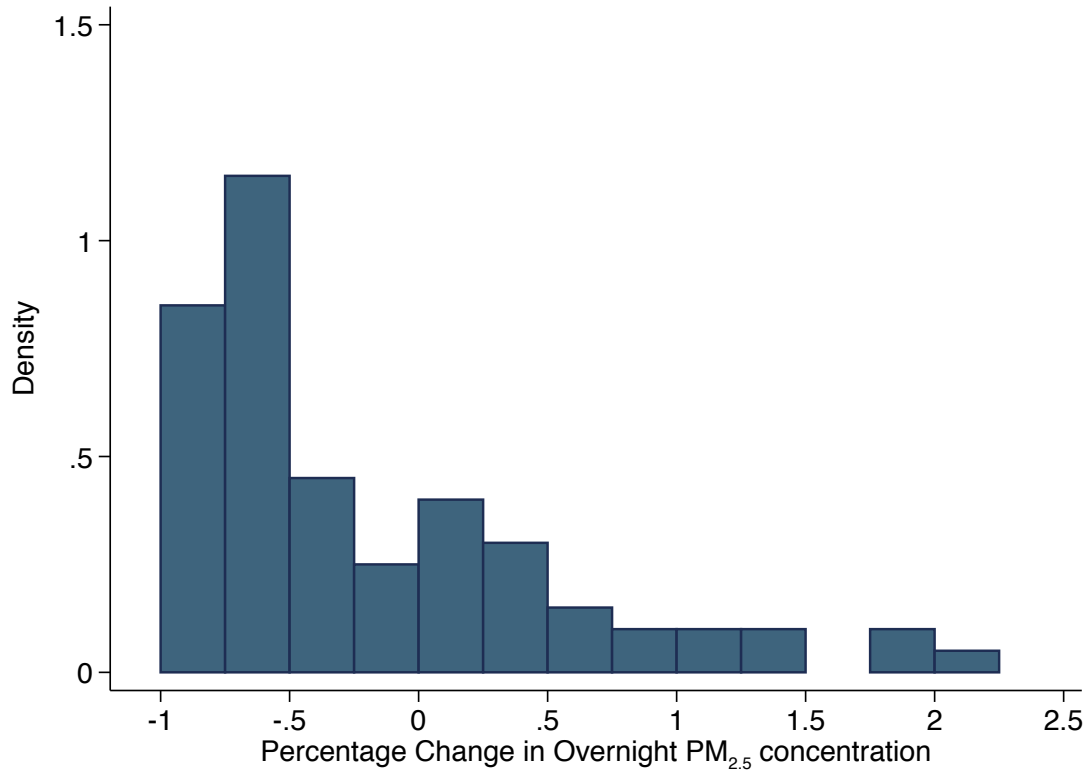
This figure shows the histogram of the unadjusted percent reductions in overnight $PM_{2.5}$ concentration (1-100%) between rounds 3 and 4 for 85 households with successful $PM_{2.5}$ measurements in both rounds, truncated at 2.5 (three households show percentage increases higher than 2.5). Forty households show reductions of 50% or higher, 14 show reductions of up to 50%, and 30 households show increases in overnight $PM_{2.5}$ concentration. Large percentage increases correspond to households with specially low levels of overnight $PM_{2.5}$ concentration at round 3. Conditional on showing an increase in $PM_{2.5}$, the average increase in levels was 0.051 mg/m^3, (95% CI 0.004 - 0.099 mg/m^3).

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
Voucher x Round 2	0.134**	0.131**	0.138**
	(0.054)	(0.053)	(0.054)
Voucher x Round 3	0.147***	0.148***	0.150***
	(0.054)	(0.054)	(0.055)
Voucher x Round 4	0.105**	0.109**	0.096**
	(0.046)	(0.046)	(0.047)
Voucher x Round 5	0.095**	0.096**	0.081*
	(0.042)	(0.043)	(0.043)
s100 x Round 2	0.123**	0.126**	0.136**
	(0.054)	(0.054)	(0.054)
s100 x Round 3	0.156***	0.156***	0.158***
	(0.049)	(0.049)	(0.049)
s100 x Round 4	0.109***	0.107***	0.099**
	(0.041)	(0.041)	(0.041)
s100 x Round 5	0.134***	0.132***	0.121***
	(0.036)	(0.036)	(0.036)
s200 x Round 2			-0.043
			(0.054)
s200 x Round 3			-0.014
			(0.048)
s200 x Round 4			0.066*
			(0.038)
s200 x Round 5			0.084**
			(0.033)
Household Fixed Effects	No	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Observations	2269	2269	2269
Number of Households	494	494	494
Mean Dep.Var.	0.56	0.56	0.56
R-squared	0.45	0.63	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. Electrification and overnight PM$_{2.5}$ concentration, experimental estimates, OLS

Panel A: Minute-Level Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	Round 3	Round 3	Round 4	Round 4	Round 5	Round 5
voucher	-1.119***	-1.316***	-0.020	0.054	0.028	-0.056
	(0.286)	(0.294)	(0.161)	(0.161)	(0.236)	(0.253)
s100	-0.340	-0.427	0.096	0.087	-0.163	-0.162
	(0.300)	(0.297)	(0.112)	(0.114)	(0.181)	(0.184)
Baseline covariates	No	Yes	No	Yes	No	Yes
Observations	86284	86284	102398	102398	106869	106869
Households	103	103	122	122	128	128
Mean Control	0.448	0.448	0.074	0.074	0.235	0.235
% Change in PM2.5	-0.674	-0.732	-0.020	0.056	0.029	-0.054
SE Change	0.093	0.079	0.158	0.170	0.243	0.239

Panel B: Household-Level Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	Round 3	Round 3	Round 4	Round 4	Round 5	Round 5
voucher	-1.119***	-1.318***	-0.171	-0.051	0.104	0.013
	(0.408)	(0.438)	(0.152)	(0.174)	(0.349)	(0.366)
s100	-0.344	-0.432	0.073	0.086	-0.135	-0.143
	(0.429)	(0.443)	(0.141)	(0.148)	(0.227)	(0.237)
Baseline covariates	No	Yes	No	Yes	No	Yes
Observations	103	103	122	122	128	128
Mean Control	0.445	0.445	0.073	0.073	0.233	0.233
% Change in PM2.5	-0.674	-0.732	-0.157	-0.050	0.110	0.013
SE Change	0.133	0.117	0.128	0.165	0.387	0.371

Notes: The dependent variable is minute-by-minute log PM$_{2.5}$ concentration from 5pm to 7am. The associated percentage change on PM$_{2.5}$, given by $e^{\hat{\beta}} - 1$, is reported in the lower panel (-1 = 100% reduction). All regressions control for hour-of-the-day, subdistrict, and monitor fixed effects. Even columns also control for baseline characteristics: sex of the household head, literacy status of the household head, use of wood for cooking and type of floor (dirt vs rest). Standard errors in parentheses, clustered at the household level. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey.*

Appendix Table. Electrification and overnight PM$_{2.5}$ concentration, IV Estimates

Panel A: Minute-Level Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	Round 3	Round 3	Round 4	Round 4	Round 5	Round 5
Connected	-3.235	-3.354	0.183	0.335	-0.215	-0.710
	(2.498)	(2.095)	(0.464)	(0.455)	(2.886)	(3.372)
Baseline covariates	No	Yes	No	Yes	No	Yes
Observations	86284	86284	102398	102398	106869	106869
Households	103	103	122	122	128	128
Mean Control	0.448	0.448	0.074	0.074	0.235	0.235
% Change in PM2.5	-0.961	-0.965	0.201	0.399	-0.194	-0.508
SE Change	0.098	0.073	0.557	0.636	2.327	1.658

Panel B: Household-Level Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	Round 3	Round 3	Round 4	Round 4	Round 5	Round 5
(mean) connected	-3.258	-3.372	-0.232	0.085	-0.954	-1.701
	(2.499)	(2.093)	(0.394)	(0.461)	(4.786)	(5.607)
Baseline covariates	No	Yes	No	Yes	No	Yes
Observations	103	103	122	122	128	128
Mean Control	0.445	0.445	0.073	0.073	0.233	0.233
% Change in PM2.5	-0.962	-0.966	-0.207	0.089	-0.615	-0.817
SE Change	0.096	0.072	0.312	0.502	1.843	1.023

Notes: The dependent variable is minute-by-minute log PM$_{2.5}$ concentration from 5pm to 7am. The excluded instruments are voucher and s100. The associated percentage change on PM$_{2.5}$, given by $e^{\hat{\beta}} - 1$, is reported in the lower panel (-1 = 100% reduction). All regressions control for hour-of-the-day, subdistrict, and monitor fixed effects. Even columns also control for baseline characteristics: sex of the household head, literacy status of the household head, use of wood for cooking and type of floor (dirt vs rest). Standard errors in parentheses, clustered at the household level. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey.*

Table 4. Electrification and overnight $PM_{2.5}$ concentration, non-experimental estimates

	(1) lpm1707	(2) lpm1707	(3) lpm1707	(4) lpm1707	(5) lpm1707	(6) lpm1707	(7) lpm1707
Connected	-0.176*	-0.168	-0.262**	-0.275***		-0.486***	
	(0.101)	(0.144)	(0.103)	(0.104)		(0.147)	
Connected x Round 3					-0.399**		
					(0.163)		
Connected x Round 4					-0.285**		
					(0.132)		
Connected x Round 5					0.131		
					(0.177)		
Connected imputed Round 4[1]							-0.328**
							(0.143)
Time FE	✓	✓	✓	✓	✓	✓	✓
Household FE		✓			✓	✓	✓
Subdistrict FE			✓	✓			
Baseline Covariates				✓			
Observations	791	791	791	783	791	599	786
Number of Households	242	242	242	240	242	240	240
Mean Control	0.21	0.21	0.21	0.21	0.21	0.21	0.21

Notes: The dependent variable is average ln($PM_{2.5}$ concentration) during the respective time window. $PM_{2.5}$ concentration was measured in the room where most household members spent most of their time awake during the evening (typically the living room). For details on the $PM_{2.5}$ measurement protocol, see appendix. Standard errors in parentheses, clustered at the household level. Col (6) restricts the sample to the first four rounds of data; Col (7) uses all the rounds but replaces connection status at round 5 with connection status at round 4. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Table 6 - Electrification and Changes in Energy Use, IV Estimates

	Kerosene		Candle		Car Battery		Propane		Wood	
	(1) RF	(2) IV	(3) RF	(4) IV	(5) RF	(6) IV	(7) RF	(8) IV	(9) RF	(10) IV
voucher_t2	-0.241** (0.112)		0.146* (0.085)		0.009 (0.046)		0.029 (0.093)		-0.054 (0.091)	
voucher_t3	-0.065 (0.099)		0.121 (0.091)		-0.006 (0.035)		-0.128 (0.134)		0.063 (0.133)	
voucher_t4	-0.091 (0.118)		0.183* (0.097)		-0.051 (0.067)		-0.068 (0.117)		-0.045 (0.105)	
voucher_t5	0.107 (0.111)		0.115 (0.122)		-0.003 (0.044)		-0.019 (0.132)		-0.047 (0.133)	
s100_t2	0.013 (0.078)		0.024 (0.062)		0.019 (0.033)		-0.077 (0.084)		0.015 (0.078)	
s100_t3	-0.005 (0.079)		0.123* (0.062)		0.013 (0.031)		-0.087 (0.086)		0.121 (0.082)	
s100_t4	-0.030 (0.077)		0.169*** (0.058)		0.048 (0.035)		-0.039 (0.084)		-0.019 (0.065)	
s100_t5	-0.071 (0.082)		0.148** (0.066)		0.032 (0.030)		-0.153* (0.087)		0.106 (0.084)	
Connected		-0.332*** (0.051)		-0.099** (0.043)		-0.017 (0.016)		0.068 (0.046)		-0.009 (0.040)
Mean Control Group										
Number of Observations	988	988	991	991	987	987	991	991	988	988

Notes: Each cell reports the IV coefficient of connection on the variable indicated in the left column at the round indicated in the top column. Regressions control for baseline characteristics: literacy, age and sex of the household head, household size, material on the walls, dirtfloor, cooking with wood, household has property title. Connection is instrumented by individual voucher allocation and the number of vouchers within 100m of the household, and controlling for the number of eligible neighbors in that radius (and the baseline characteristics, as in equation 1). Huber-White standard errors in parentheses. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Table 7 - Acute Respiratory Infections Among Children 0-6, LPM

	(1) RF	(2) IV	(3) RF	(4) IV
Voucher x Round 3	-0.180*		-0.160*	
	(0.096)		(0.097)	
Voucher x Round 4	0.091		0.106	
	(0.070)		(0.073)	
Voucher x Round 5	-0.032		-0.011	
	(0.089)		(0.091)	
s100 x Round 3	-0.044		-0.048	
	(0.088)		(0.088)	
s100 x Round 4	-0.006		-0.032	
	(0.087)		(0.085)	
s100 x Round 5	-0.055		-0.071	
	(0.076)		(0.080)	
Connected		-0.652*		-0.660*
		(0.389)		(0.359)
Baseline Characteristics	No	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Subdistrict FE	Yes	Yes	Yes	Yes
Mean Control, Round 3	0.43	0.43	0.43	0.43
Mean Control, Round 4	0.10	0.10	0.10	0.10
Mean Control, Round 5	0.23	0.23	0.23	0.23
Observations	546	546	538	538
Individuals	192	192	189	189

Notes: The dependent variable in all regressions is an indicator for acute respiratory infection in the four weeks prior to the survey (self-reported). Baseline characteristics: literacy, age and sex of the household head, household size, material on the walls, dirtfoor, cooking with wood, household has property title. Excluded instruments in columns (2) and (4) are voucher and s100 interacted with round. Standard errors, clustered at the household level, are reported in parenthesis. Significantly different than zero at 90(*), 95(**), and 99(***) percent confidence. Source: *Household Electrification Survey*

Table 8 - Estimation of PM$_{2.5}$ Exposure and Health Impacts

	Female Head	Male Head	Female Child	Male child
Time (hours per day)				
Sleeping	8.85	8.83	9.99	9.54
Kitchen	2.54	0.08	0.56	0.09
Outside	2.69	7.77	2.51	4.88
Home, Evening	5.15	5.17	4.01	4.46
Home, Daytime	4.77	2.14	6.93	5.03
Air inhalation (m^3) per hour				
Sleeping	0.30	0.30	0.30	0.30
Light Activity	0.78	0.78	0.78	0.78
Estimated exposure to PM$_{2.5}$ (mg per day)				
Non-encouraged group	5.68	3.20	4.23	3.72
Encouraged group	3.93	1.43	2.65	2.10
% Change in exposure due to electrification (lower bound)				
Percentage change	-0.328	-0.587	-0.392	-0.457
Bootstrapped standard errors	0.001	0.001	0.003	0.003
Predicted relative risk				
Lung cancer				
Non-encouraged group	3.98	2.68	3.22	2.95
Encouraged group	3.06	1.75	2.39	2.10
Ischemic heart disease				
Non-encouraged group	1.12	1.07	1.09	1.08
Encouraged group	1.08	1.03	1.06	1.05
cardiovascular disease				
Non-encouraged group	1.12	1.07	1.09	1.08
Encouraged group	1.08	1.03	1.06	1.04
Cardiopulmonary disease				
Non-encouraged group	1.15	1.08	1.11	1.10
Encouraged group	1.10	1.04	1.07	1.06

Notes: Exposure $= \sum_{j=1}^{J} time_j \times concentration_j \times inhalation_j$, where $time_j$ is time spent in activity j, *concentration* is PM$_{2.5}$ concentration in the room while performing activity j, and $inhalation_j$ is the esimated inhalation rate while performing activity j. Time allocation was obtained from the *EHEIPCER* time allocation module. "Home, daytime" is time at home between 0700-1700hrs in any room but the kitchen. "Home, evening" is the time at home between 1700-0700 hrs, while neither sleeping nor in the kitchen. "outside" is time spent outside the home. Air inhalation was taken from the EPA exposure handbook (EPA, 2011). "Light activities" include watching TV, desk work, writing, typing, cooking, washing dishes, ironing, walking up to 2.5 mph (2.9 km/h). PM$_{2.5}$ concentration was estimated at 0.41 mg/m^3 for "sleeping" and "home, evening", 0.26 mg/m^3 for "home, daytime", 0.90 mg/m^3 for "kitchen", 0 mg/m^3 for "outside". For the "after electrification" scenario, PM$_{2.5}$ was reduced to 0.15 mg/m^3 (by 63%, as estimated by our main model) for "sleeping" and "home, evening" and held constant for the other instances. Average ages: 43 (female head), 47 (male head), 11 (female child), 13 (male child). Predicted relative risks were calculated with a linearized version of the dose-response function calculated by Pope III et al. (2011). R^2 for lung cancer (0.99), ischemic heart disease (0.96), cardiovascular disease (0.86), cardiopulmonary disease (0.80). See section 6 for details.

Appendix 1 - Summary Statistics by Subsample

	Not connected	Voucher Subsample	Air Quality Subsample Experimental	Non-Exp.
Age of household head	49.02	49.54	48.18	50.79
	(17.55)	(16.82)	(17.49)	(17.32)
Household head is male	0.68	0.69	0.73	0.73
	(0.47)	(0.46)	(0.44)	(0.45)
Average age in hosehold	31.23	30.80	29.84	33.25
	(17.72)	(17.05)	(16.43)	(19.04)
Household size	4.34	4.53	4.64	4.53
	(2.36)	(2.32)	(2.16)	(2.50)
Total dependency ratio	0.44	0.45	0.42	0.44
	(0.28)	(0.27)	(0.26)	(0.28)
Maximum schooling in the household	5.47	6.05	5.90	4.21
	(4.35)	(4.34)	(4.23)	(3.13)
Schooling of the household head	2.23	2.36	2.53	1.45
	(3.55)	(3.47)	(3.71)	(2.18)
household head is literate	0.48	0.54	0.57	0.39
	(0.50)	(0.50)	(0.50)	(0.49)
Income pc, 1000USD per year	0.65	0.77	0.62	0.43
	(1.72)	(4.01)	(1.11)	(0.67)
Monthly expenditure in kerosene	2.49	2.11	2.13	4.83
	(4.17)	(3.90)	(3.65)	(5.12)
Monthly expenditure in propane	1.82	2.09	1.96	0.88
	(2.86)	(2.96)	(3.15)	(2.25)
Monthly expenditure in candles	0.50	0.46	0.36	0.80
	(1.65)	(1.57)	(1.32)	(1.91)
Monhtly expenditure in car battery rchg	0.12	0.08	0.12	0.25
	(0.65)	(0.56)	(0.65)	(0.92)
Cooks with wood	0.72	0.70	0.70	0.90
	(0.45)	(0.46)	(0.46)	(0.30)
Informal electricity	0.46	0.38	0.46	0.04
	(0.50)	(0.49)	(0.50)	(0.20)
Agrees with the following statement				
Electricity illuminates better than kerosene.	0.95	0.96	0.96	0.95
	(0.22)	(0.20)	(0.20)	(0.22)
Powering a TV is cheaper w/elect than battery.	0.76	0.79	0.81	0.77
	(0.43)	(0.41)	(0.39)	(0.43)
Cooking with electricity is not convenient	0.58	0.52	0.53	0.53
	(0.49)	(0.50)	(0.50)	(0.50)
Electricity is very expensive	0.52	0.49	0.54	0.40
	(0.50)	(0.50)	(0.50)	(0.49)
Woodsmoke generates respiratory problems	0.88	0.86	0.86	0.88
	(0.32)	(0.35)	(0.35)	(0.33)
Kerosene is not an expensive source of lighting	0.34	0.35	0.41	0.34
	(0.48)	(0.48)	(0.49)	(0.48)
Kerosene is the best way to illuminate my household	0.22	0.21	0.25	0.28
	(0.41)	(0.41)	(0.43)	(0.45)

Notes: Samples - Col(1): all *EHEIPCER* households that were off-grid at baseline (N=2014). Col(2): the households among whom vouchers were randomly allocated (N=500). Col(3): random subset of the households in column 2 (N=150). Col(4): *EHEIPCER* households that had not connected to the grid by the first follow-up (N=207), and from the same departments (roughly equivalent to counties in the US) as those in column2. See main text for details.