NBER WORKING PAPER SERIES

UP IN SMOKE: THE INFLUENCE OF HOUSEHOLD BEHAVIOR ON THE LONG-RUN
IMPACT OF IMPROVED COOKING STOVES

Rema Hanna
Esther Duflo
Michael Greenstone

Working Paper 18033
http://www.nber.org/papers/w18033

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2012

Up in Smoke: The Influence of Household Behavior on the Long-Run Impact of Improved
Cooking Stoves
Rema Hanna, Esther Duflo, and Michael Greenstone
NBER Working Paper No. 18033
May 2012
JEL No. I15,I18,O10,O12,O13,Q0,Q23,Q3,Q51,Q53,Q56

ABSTRACT

It is conventional wisdom that it is possible to reduce exposure to indoor air pollution, improve health outcomes, and decrease greenhouse gas emissions in the rural areas of developing countries through the adoption of improved cooking stoves. This belief is largely supported by observational field studies and engineering or laboratory experiments. However, we provide new evidence, from a randomized control trial conducted in rural Orissa, India (one of the poorest places in India), on the benefits of a commonly used improved stove that laboratory tests showed to reduce indoor air pollution and require less fuel. We track households for up to four years after they received the stove. While we find a meaningful reduction in smoke inhalation in the first year, there is no effect over longer time horizons. We find no evidence of improvements in lung functioning or health and there is no change in fuel consumption (and presumably greenhouse gas emissions). The difference between the laboratory and field findings appear to result from households' revealed low valuation of the stoves.   Households failed to use the stoves regularly or appropriately, did not make the necessary investments to maintain them properly, and usage rates ultimately declined further over time. More broadly, this study underscores the need to test environmental and health technologies in real-world settings where behavior may temper impacts, and to test them over a long enough horizon to understand how this behavioral effect evolves over time.

Rema Hanna
Kennedy School of Government
Harvard University
79 JFK Street
Cambridge, MA 02138
and NBER
Rema_Hanna@hks.harvard.edu

Esther Duflo
Department of Economics
MIT, E52-252G
50 Memorial Drive
Cambridge, MA 02142
and NBER
eduflo@mit.edu

Michael Greenstone
MIT Department of Economics
50 Memorial Drive, E52-359
Cambridge, MA 02142-1347
and NBER
mgreenst@mit.edu

I. INTRODUCTION

Half of the world's population, and up to 95 percent in poor countries, rely on solid fuels, including biomass fuels (e.g., wood, dung, agricultural residues) and coal, to meet their energy needs. The World Health Organization lists "indoor air pollution (IAP) from primitive household cooking fires as the leading environmental cause of death in the world," stating that "it contributes to nearly 2.0 million deaths annually," about as many deaths as malaria and tuberculosis combined (Martin, Glass, Balbus, and Collins, 2011). Moreover, cooking with biomass fuels is a key source of climate change through its releases of carbon dioxide (CO_2) and black carbon (Kandlikar, Reynolds, and Grieshop 2009). In response, improved cooking stoves are increasingly seen as a tool to improve respiratory health and combat climate change.[1] For example, in September 2010, Hillary Clinton announced the formation of the Global Alliance for Clean Cookstoves (GACC), which calls for 100 million homes to adopt clean and efficient stoves and fuels by 2020. However, this big push for improved cooking stoves has occurred despite surprisingly little rigorous evidence on their efficacy on health and fuel use.[2]

This paper reports the results from a randomized evaluation of improved cooking stoves on individuals' behavior and well-being. Specifically, we evaluate the effect of distributing an inexpensive, improved stove on health and fuel use in Orissa, India. Gram Vikas (GV), an award winning nongovernmental organizations (NGO), obtained funding to subsidize stove construction for 15,000 households over five years, independent of the research.[3] GV chose stoves designed with an enclosed cooking chamber (to keep the flame separate from the food) and a chimney to direct smoke away from the user. The stoves had been proven to reduce IAP and energy consumption in laboratory settings and could be constructed with locally sourced materials, facilitating distribution at a large scale. At a total cost of about US$12.50, these stoves fall within the "lower end" of improved stove technologies. However, these stoves represent the vast majority of improved stoves that have been distributed: the World Bank (2011) reports that

---

[1] These benefits are cited regularly in leading publications, including *The New York Times*, and they have a range of proponents from Bill Clinton to Julia Roberts.

[2] According to the GACC website, the stoves (1) reduce child pneumonia by 50 percent, (2) save the equivalent of 1–2 tons of CO_2 per year, and (3) produce fuel savings that families can use to pay for the stove. However, as we discuss, none of the evidence, to date, fully supports these claims.

[3] Gram Vikas has won numerous awards, including being listed in the *Global Journal*'s "Top 100 Best NGOs in the World" in 2012, and has attracted considerable international funding.

stove programs have typically distributed improved stoves in this category and over 166 million of them are in use today.

We used a public lottery to randomly assign the order in which stoves were constructed within each village for 2,600 households that were part of GV's stove program. The first third of households within each village received the stoves at the start of the project, the second third received the stoves about two years after the first wave, and the remaining households received it at the end of the study. Households were followed for four years after the initial stove offers, which allows for an examination of the long-run use and impacts of the stoves. This long-run follow-up is virtually unprecedented in evaluations of health interventions or other new technologies where households learn about the benefits and maintenance needs over time.

There are four primary results. First, initial household take-up and usage of the (almost free) new stoves was far from universal and then declined markedly over time as households failed to make the maintenance investments (e.g., cleaning the chimney) necessary to keep them fully operational. Several measures document this, but perhaps the most salient is that treatment households that received the GV improved stoves still continued to use their traditional stoves in conjunction with the new ones—even early on, when the majority of the stoves were functional. In the early years, treatment households only cooked 3.5 more meals per week (or 25 percent of total meals) with a good condition, improved stove than the control households.[4] This difference was halved to about 1.8 meals per week in year 3, as the stoves deteriorated.

Second and correspondingly, the stoves failed to achieve their primary goal of reducing exposure to hazardous air pollutants. While there was a significant effect on smoke inhalation during the first year for the primary cooks in the household, the treatment effect became statistically indistinguishable from zero in subsequent years as proper stove usage declined. Further, even in the first year, the resulting effect (a 7.5 percent decrease in the carbon monoxide concentration of exhaled breath) was smaller than the reduction observed in laboratory-style settings with properly maintained stoves and near perfect usage rates.[5]

---

[4] Levine and Beltramo (2010) observed the same phenomenon with solar ovens (another type of stove used to reduce smoke exposure and energy consumption) in Senegal: even households that chose to use the solar oven generally cooked only a few of their meals on them, continuing to cook the remaining meals on their standard stoves.

[5] We were also unable to detect an effect on females more generally or on children who were frequently near stoves during cooking.

Third, we cannot reject the null hypothesis that the stoves failed to affect health across a wide set of health outcomes. For example, there is no difference in lung functioning (as measured by spirometry tests) between women who regularly cook in the treatment and control groups. Furthermore, we fail to find a positive impact on a wide variety of measured and self-reported health outcomes, including infant birth weight, infant mortality rates, probability of a cough, blood pressure, or even the probability of any illness in the last 30 days.

Fourth, the treatment group appears to have experienced modest declines in their living standards and there is no evidence of a reduction in greenhouse gas emissions. Specifically, treatment households spent substantially more time repairing their stoves. Furthermore, the treatment did not affect fuel costs or time spent cooking, which is consistent with the energy consumption results of "health-improving" stoves studied by Miller and Mobarak (2011) and Burwen and Levine (2011). There is also no evidence of potential climate benefits from reductions in deforestation since there is no change in total wood used for cooking. It is noteworthy that these findings contrast self-reported satisfaction of improved stoves and laboratory test results that show reduced time and energy used to boil the same quantity of water with an improved stove.

Besides demonstrating the importance of human behavior in assessing the effectiveness of new technologies, this study builds upon and contributes to the literature on indoor air pollution. Most evidence on IAP comes from observational studies that compare fuel use and health status of users and nonusers (Bruce, Perez-Padilla, and Albalak, 2000). However, households that cook with improved stoves are typically different in other respects as well, such as income levels and health preferences (Bruce, Neufeld, Boy, and West, 1998; Mueller, Pfaff, Peabody, Liu and Smith, 2011). Thus, despite the positive effects of reducing indoor air pollution in this literature, it is unclear whether these results reflect the impact of improved stoves or unobservable characteristics.

Recently, more rigorous evidence has emerged from the RESPIRE study, a randomized experiment of a concrete stove in Guatemala (Smith-Sivertsen, Diaz, Pope, Lie, Diaz, McCracken, Bakke, Arana, Smith, and Bruce, 2009). Our paper complements this study in at least two important ways. First, we followed households for four years after the receipt of the stoves, compared to the RESPIRE's follow-up of 12 months for the full sample and 18 months for a subset, to measure long-run impacts. Our extended evaluation may be important for at least

several reasons. First, the treatment effects on health may change considerably over time, as households learn about the value of the stoves and subsequently change their usage rates and maintenance investments, as well as experience a general depreciation of the technology.[6] Second, the effect on health may be cumulative over several years. Third, we found meaningful effects on CO for primary cooks in the first year; had we ended the study after learning this, we would have projected the effect for several years in benefit-cost calculations. In reality, this effect was short-lived.[7]

Second, we study an actual program run by a local non-profit with no assistance by the research team. The stoves are locally made and relatively cheap (roughly $12.50), implying that they would be practical for large scale distribution and presumably affordable for the target population if sold (annual per-capita consumption of households in our sample is $145). The RESPIRE stoves cost between $100 and $150, which makes them prohibitively expensive for most households where indoor air pollution is a problem. Furthermore, although the RESPIRE study is conducted in the field, trained fieldworkers inspected the stoves weekly for proper use and maintenance and then arranged for repairs as needed (Smith, McCracken, Thompson, Edwards, Shields, Canuz and Bruce, 2010). In this respect, the RESPIRE study shuts down households' ability to reveal their valuation through usage rates and decisions about shifting resources from other goods to stove maintenance. Thus, the results from the RESPIRE study likely provide upper bound effects, while our estimates more closely resemble real-world impacts, where households may not use the technology appropriately or may choose not to use the technology at all. The mixed results on health from the RESPIRE study (discussed in depth below) and the lack of health impacts found in our study, which derive from limited and improper use, suggests that, in the context of evolving stove technologies, the new generation of stoves (e.g. envirofit and rocket stoves) need to be evaluated in field settings to understand if they will actually provide the benefits measured in the laboratory before valuable resources are devoted to their deployment.

---

[6] For example, Dupas (2012) shows that a one-time subsidy can boost the purchase of malaria bednets because it gives individuals a chance to learn about the technology and positively update their beliefs. However, take-up and use of the stoves may also decline over time, if households learn that the product does not perform to their expectations, either because the technology does not provide the promised benefits or the signal on the benefits is noisy.

[7] Note that our study's sample included over 4,000 women and 3,000 children, compared to the 500 women and children in the RESPIRE study, which provides us with greater precision in detecting any health and fuel effects.

More generally, this paper contributes to the literature on the adoption of health and environmental technologies. Many times, new technologies are evaluated in laboratory experiments or through field experiments in controlled settings where researchers ensure high compliance in terms of use and maintenance. These studies are vital because they provide an upper bound effect on the possible treatment effects of the technologies. However, as Chassang, Padro-i-Miguel, and Swoberg (2011) discuss, perfectly controlling individual choices and actions produces an impact estimate that is internally valid, but may lack external validity to the large scale distribution of the technology, where field implementation may be less than ideal. This is especially true for health and environmental technologies whose benefits often stem from proper and sustained use.[8] This study demonstrates the additional value of evaluating a technology in real-world settings and for long enough to understand how individuals' behavioral responses influence the technology's effectiveness.

The remainder of the paper proceeds as follows. Section II discusses the experimental design. The data is described in Section III, while the empirical framework is laid out in Section IV. The findings are presented in Section V. Section VI provides a discussion of the state of knowledge on improved stoves to date, as well as lessons that can be learned for future research and evaluation. Section VII concludes the paper.

II. EXPERIMENTAL DESIGN

A. Setting

This project took place in India, where about 70 percent of the population burn solid fuels—firewood, crop residue, or cow dung—in traditional stoves (see Appendix Figure 1, Panel A) to meet their cooking needs (Census of India, 2001). The reliance on traditional fuels is even higher (90 percent) in poorer, rural regions. Indoor air pollution (IAP) levels from traditional

---

[8] In Indonesia, Thomas et al. (2006) test the effect of iron supplementation on work outcomes, where compliance is carefully monitored. They find large effects on work, providing an upper bound on the effect in large scale distributions where individuals may not regularly take the supplements. Similarly, key randomized evaluations of bed nets (Alonso et al., 1991; Phillips-Howard et al., 2003; Binka, 1998; Nevill et al., 1996) typically send project staff to re-treat the nets every six months, where presumably households are also reminded to use them. In Phillips-Howard et al. (2003), households signed forms that the nets remained the property of the project until after the study was concluded, which could have induced households to keep the nets in better condition than if they had been told that the nets were their own. This issue extends beyond the developing world: Duggan (2005) argues that in the United States anti-psychotic drugs are less effective in practice than in FDA trials due to several factors, including short-run follow-up, small sample sizes that make it difficult to detect some important side effects, and prescribing these drugs for people that differ from the individuals enrolled in the clinical trials.

stoves are high. For example, Smith (2000) reports that the "available data show a distribution of indoor PM_{10} 24-h concentrations measured in Indian solid-fuel-using households ranging to well over 2000 µg/m^3." To put these figures into context, the Central Pollution Control Board of India states that ambient levels of PM_{10} should not exceed 100 µg/m^3.

In response to the health threats from solid fuel use in traditional stoves, as well as concerns about deforestation, both governments and nongovernmental organizations (NGOs) have been implementing clean stove programs for several decades. For example, during the 1980s and 1990s, the Indian government alone subsidized and distributed 32 million improved stoves. However, many of these stoves had life spans that were less than two years, and as Smith (2000) has pointed out, only a small fraction of the stoves built before 1990 still existed at the time of his article. In fact, this campaign is widely acknowledged to have been a failure, with stoves laying unused or rapidly falling into disrepair (Block 2011).[9] The renewed interest in IAP worldwide has prompted a new wave of interest in India as well, with NGOs, local governments, and private foundations investing in the design and distribution of improved stoves, and the launch in 2011 of a new large-scale government program with an improved design.

This paper evaluates an improved stove program run by Gram Vikas, an NGO that operates in the state of Orissa. Orissa is one of the poorest states in India, with 40 percent of the population living below the poverty line. Poverty is significantly worse in the western and southern districts of the state where this project took place. Independent of the researchers, Gram Vikas obtained funding from the Inter-Community Church Organization (ICCO) to subsidize the construction of improved stoves to roughly 15,000 households over five years.

The stove considered in this study is a relatively inexpensive improved stove technology. It was developed and tested by the Appropriate Rural Technology Institute (ARTI), an NGO specializing in energy innovation for rural areas. Like the traditional stoves, it is largely made out of mud (see Appendix Figure 1, Panel B). However, the constructed base encloses the cooking flame and it includes a chimney to direct smoke away from the user. Moreover, it allows for two pots, instead of the one pot in traditional stoves, to potentially reduce cooking time.

---

[9] Rita Colwell, the former director of the US National Science Foundation, is cited in this article as saying "You can't drop a stove into a household and walk away. You need to do follow-up. You need implementation."

The chosen stove was considered appealing because it is constructed with local materials and is low cost, at roughly US$12.50. Gram Vikas subsidized the stove cost by contributing stove materials (chimney), design, and a skilled mason to supervise the construction. Households were responsible for providing mud for the stove base, labor and a payment of about US$0.75, which was used to pay the mason who assisted in building and maintaining the stoves, as well as a fund for new stoves if a new house is built in the village. As the stove is made from locally available materials, it can be easily constructed in these remote, rural areas of India.

In laboratory settings, the ARTI stove burns more efficiently than a traditional stove, leading to lower biofuel requirements and less indoor smoke. However, obtaining this outcome requires that the stoves are maintained appropriately, which involves repairing cracks and regularly removing chimney obstructions. Moreover, households must place the pots on the openings correctly, and cover the second pot when it is not in use in order to avoid smoke from escaping.

In addition to providing the stoves, Gram Vikas conducted the standard information campaigns that NGOs run when they introduce a new program. Specifically, during construction they held training sessions on proper use and maintenance (see Appendix Figure 2 for an example of the training materials). Among households that received a stove in the first wave, almost 70 percent report that they attended a training session. Moreover, Gram Vikas identified individuals in each village who used their stoves correctly and hired (with a small stipend) them to promote proper use and alert Gram Vikas when the stoves were in need of repair. Of those who received a stove in the first wave, 62 percent report knowing who this "promoter" is, 48 percent report that they attend a meeting with the promoter, and another 47 percent state that they received a visit from the promoter to discuss stove use. In total, about 86 percent report either having Gram Vikas or the promoter provide training on the stove (either through a meeting or visit).

B. Sample, Timeline and Experimental Design

In the summer of 2005, Gram Vikas obtained permission from 42 villages to participate in the study. Unrelated to the study, three villages decided to withdraw from all Gram Vikas activity.[10]

---

[10] Gram Vikas's main activity is the construction of village sanitation systems (a toilet and a tap for each house). This requires the cooperation of the entire village (across caste lines). This occasionally leads to the breakdown of

As a result, we added five additional villages in June 2007. Therefore, a total of 2,651 households in 44 villages participated in the study.

After we completed the baseline survey in each village (in 2006 for the majority of villages, and in 2007 for the additional five villages; see timeline in Appendix Figure 3), a village meeting was conducted. At each meeting, Gram Vikas explained that the stoves were being built in three waves, and that the households would be randomly assigned to each wave. Next, a public lottery (monitored by the research team) was conducted to choose the first third of households in the village that would be offered a GV improved stove. Gram Vikas completed the first wave of stove construction and user training between September 2006 and March 2007.

After we conducted the midline survey, the second round of village meetings occurred. A lottery was conducted to choose households that would be offered a stove in the second wave of construction. From May 2009 to April 2010, the second round of stove construction and training occurred. Note that during this time, there was also a big push by Gram Vikas to repair or rebuild stoves from the first wave of construction.

III. DATA

A. Data Collection

Throughout the study, we conducted a series of surveys to create a panel dataset on stove use, smoke exposure, health, stove breakages and repairs, and fuel use. Appendix Figure 4 provides a summary of the surveys and their sample sizes, while the data appendix describes each survey that we conducted in more detail. Here, we provide a summary of the key variables of interest.

We collected comprehensive data on the socio-demographic characteristics of each household. This data includes household composition (size, as well as each member's age, sex, and relationship to the head of household), demographics (education levels, caste, religion), economic indicators (assets, indebtedness), and consumption patterns. In addition, for each household member, we collected individual measures of productivity, such as employment status, time-use patterns for adults over the last 24 hours, and school enrollment and attendance for children.

Through a series of surveys, we collected information on stove use. This included the

---

their relationship with villages. See Duflo, Greenstone, and Guiteras (2011) on the benefits of their sanitation program.

types of stoves a household owned, meals cooked with each type of stove over the past week, repairs and maintenance activities surrounding the stoves, and fuel expenditures (both money and time). In addition, we collected information on beliefs about the efficacy of the stoves (for example, do they use less fuel) and on satisfaction with the stoves.

To measure smoke exposure, the team measured exhaled carbon monoxide (CO) with a Micro Medical CO monitor.[11] CO is a biomarker of recent exposure to air pollution from biomass combustion, and therefore it can be used to proxy an individual's personal exposure to smoke from their stoves. Furthermore, it is an inexpensive way to proxy for inhalation of particulate matter, which has been shown to be an important determinant of infant mortality and life expectancy (see, for example, Chay and Greenstone 2003a, 2003b; Chen et al., 2012; Currie and Neidell, 2005; Jayachandran, 2009; Arceo-Gomez, Hanna, and Oliva, 2012).

Finally, we collected two types of health data. First, we conducted detailed health recall surveys where we enquired about symptoms (coughs, colds, etc.), infant outcomes, and health expenditures. We complemented these data with physical health checks for biometric measurements, such as height, weight, and arm circumference. During the physical health check, we administered spirometry tests that are designed to gauge respiratory health by measuring how much air the lungs can hold and how well the respiratory system can move air in and out of the lungs. In contrast to peak flow tests, which are easier to administer, spirometry readings can be used to diagnose obstructive lung disorders (such as chronic obstructive pulmonary disease [COPD] and asthma), and also restrictive lung disorders.[12] Further, this test is the only way to obtain measurements of lung function that are comparable across individuals (Beers et al., 1999). The tests were conducted using the equipment directions, as well as guidelines from the

---

[11] Note that we did not measure ambient pollutants (neither CO nor PM). Ambient measures alone are less interesting to measure than exposure, as individuals may undertake fewer behaviors to protect themselves from smoke if ambient measures fall and thus could, in fact, end up with a higher level of exposure. If we conducted only ambient measures we could see a decline, even though their actual exposure may not have decreased due to these behavioral changes. We focused on CO, which has been argued to be a good proxy for PM. Collecting data on PM exposure is difficult in this setting: tubes must be attached to the subjects for 24 hours and the equipment requires controlled temperature, careful transferring of samples, and proper laboratories for testing. Given the conditions of rural Orissa, controlling the samples would be near impossible on such a large scale. However, McCracken and Smith (1998) report a strong correlation between the average concentrations of CO and $PM_{2.5}$ in the kitchen during water boiling tests. They conclude that this implies "the usefulness of CO measurements as an inexpensive way of estimating $PM_{2.5}$ concentrations," even if it is not an exact proxy (see Ezzati (2002) for a discussion of this).

[12] According to the Global Initiative for Chronic Obstructive Lung Disease, the results can be used to assess whether participants have chronic obstructive pulmonary disease (COPD). There are two main forms of COPD, chronic bronchitis and emphysema; complications from the disease include heart failure, pneumonia, severe weight loss, and malnutrition.

American Association for Respiratory Care.[13]

Lastly, throughout the study, we compiled Gram Vikas's administrative data on program functioning. Specifically, we collected data on lottery participation and outcomes.

B. Sample Statistics

Table 1 provides information on household-level baseline demographic characteristics and stove usage. For each variable, means are provided in Column 1, standard deviations in Column 2, and the sample sizes in Column 3. As Panel A indicates, the households were very poor, with an average monthly per capita household expenditure of about US$12 (Rs 475). Forty-three percent of households belonged to a disadvantaged minority group. A little less than half had electricity, making electric stoves an impractical option. Schooling outcomes are discouraging: only 69 percent of the household heads had attended school, and just 58 percent self-reported being able to read. Similarly, only 32 percent of the female household heads (or spouses of one) had attended school, with just 20 percent self-reporting that they are literate.

There was a large dependence on traditional stoves and fuels for cooking (Panel B). Most households (99 percent) owned at least one traditional cooking stove (see Appendix Figure 1, Panel A). About a quarter of households owned any type of low polluting stove, primarily electric (11 percent) and kerosene (10 percent). Only 23 out of 2,480 households had an improved stove from a previous program that Gram Vikas had conducted several years earlier. Despite the fact that many households owned a low-polluting stove, most (93 percent) continued to use the traditional stove as their primary stove, with an average of 12.6 meals (or 91 percent of all meals) cooked on one over the last week. Unsurprisingly, out of those who primarily cook household meals, 96 percent were female.

Given that households often had more than one stove, they tended to cook in more than one location (Panel C). On average, about one meal per week was cooked in an enclosed area, and about five meals per week were cooked in a semi-enclosed area. About 7.5 meals per week were cooked outside. It is noteworthy that open fire stoves pollute enough that they produce significant exposure even when used outdoors (Smith, 2000).

---

[13] A manual spirometer was used in the baseline, continuous health survey, and a portion of the midline. The enumerators would take up to seven readings for each individual, until there were at least three satisfactory readings and at least two FEV1 readings within 100mL or 5 percent of each other. Electronic spirometers were adopted halfway through the midline. The new machines indicated when satisfactory readings had been completed and saved the best reading for each individual.

Households relied heavily on wood for fuel (Panel D), with 99 percent reporting having ever used wood as fuel. On average, about 5kg of wood was used to cook the household's last meal. Fuel was typically obtained by a combination of collection and purchases: 83 percent report having ever gathered wood, and 35 percent report having ever bought wood for cooking. About 20 percent of households also report having ever sold wood. About 10 percent report having collected wood the previous day, spending about five hours, on average, if they did so.

Table 2 presents baseline health statistics. Panel A reports on adults (all women in Columns 1–3 and households members who identified themselves as primary cooks in the baseline in Column 4–6), while Panel B reports on children (all children in Columns 1–3, and children aged 5 and under in Columns 4–6).

Women had high levels of smoke exposure: the average CO reading is 7.55 ppm, where a reading between 6–9 ppm indicates smoke within the lungs and a reading of 10 ppm or more indicates a high level of smoke within the lungs.[14] About 27 percent of women score a reading of 10 ppm or more, which, following the back-of-the-envelope calculation by Levine and Beltramo (2010), suggests that these women had exposure levels that were equivalent to smoking 10 cigarettes per day (note that few women reported that they smoked). In contrast, lung function measurements are in the normal range. We observe a mean FEV1/FVC of about 90, which suggests that, on average, participants did not have COPD.

In general, self-reported illness levels among women were high. Almost 90 percent report having had any type of symptoms in the past 30 days. Symptoms that are typically associated with smoke exposure were abundant: about half self-reported having had a cough or cold in the last 30 days, about 49 percent report having had a headache, and about 30 percent report having had sore eyes. In contrast, very few individuals report that they experienced tightness in their chest (5 percent) or wheezing (2 percent). Relatively speaking, health expenditures were high, with females reporting that they spent about US$1.63 in the last month.[15] Note that the baseline health statistics for household primary cooks are similar to those for the sample of adult women.

---

[14] The baseline CO in exhaled breath is slightly lower than in the RESPIRE study, which found a baseline CO rate of about 9 ppm (Diaz et al., 2007). However, it is similar to the control group mean in the RESPIRE study of about 7 ppm that was observed throughout the course of that study (Smith-Sivertsen et al., 2009).
[15] Banerjee, Duflo, and Deaton (2004) find similar health results in Rajasthan, India.

Children had high levels of CO exposure and poor health outcomes in the baseline (Panel B). Their CO levels were, on average, 6.48 ppm.[16] This suggests that they had an average CO level similar to someone who smokes about 7 cigarettes per day. About 20 percent of the children had a reading of 10 ppm or higher, which is equivalent to being a heavy smoker. Children were malnourished, with an average BMI nearly two standard deviations below the norm, according to the 2000 US Centers for Disease Control measurement of the child population (two standard deviations below the norm is generally considered an indicator of stunting). Parents report that 73 percent of the children had some form of illness in the past month. About a quarter of parents consulted a health care provider for a child's fever in the last month, with an average of US$1.15 spent on all healthcare costs during this period. Coughs were the most prevalent symptom, with about 40 percent of all children having had one in the last 30 days. Other illnesses that could be associated with indoor air pollution include ear infections (9 percent), skin irritation (13 percent), and vision problems (1 percent). In general, children 5 years of age and younger (Columns 4 to 6 of Panel B) look similar in terms of baseline characteristics to the sample of all children.

C. Is Cooking on a Clean Stove Associated with Better Health? Baseline Correlations

Most of the evidence on health improvements from reductions in indoor air pollution is based on the association between clean stove usage and health in observational data. In Table 3, we present the coefficient estimates from regressions of baseline CO and health variables on the number of meals cooked with a low polluting stove in the last week at baseline, conditional on village fixed effects and several demographic variables (indicators for whether households belong to a disadvantaged minority; the household has electricity; the male household head has attended school; the male household head is literate; the female household head (or wife of male household head) has attended any school and is literate; any female member within a household has a savings account; and monthly household per capita consumption).[17] Panel A presents

---

[16] Note that in the baseline and CHS, only children approximately 9 years and older were tested for CO exposure, as it is difficult to test younger children. Based on a doctor's assessment and field testing, we lowered the age restriction and collected CO measures for children older than 5 years in the midline and endline.

[17] Appendix Table 1 presents these estimates unconditional on any control variables. In these specifications, we observe that meals are also positively correlated with health expenditures and negatively correlated with wheezing for women and primary cooks. In general, the coefficient estimates for women and primary cooks tends to be a bit larger in magnitude and significance when we do not include the control variables, but they are not qualitatively different from those presented in Table 3. The estimates for children are also similar to those presented in Table 3.

these results for females, while Panel B reports them for the primary cooks.   Panels C and D present results for all children and those aged five and under, respectively. The standard errors in all regressions are clustered at the household level.

Consistent with the literature, cooking with a clean stove is associated with better health. As shown in Panel A, women who cooked more meals with any type of low-polluting stove had better lung functioning (Column 3), higher BMI (Column 4), fewer coughs or colds in the last week (Column 5), and a lower likelihood of sore eyes (Column 6) and wheezing (Column 9). The coefficient estimates are similar for the primary cooks (Panel B). Note that while we observe that cooking an additional meal with an improved stove is associated with lower smoke exposure, as measured by CO for both women and primary cooks (Column 1), this is not significant at conventional levels.

For the children who were tested in the baseline for CO exposure, there is a significantly negative correlation between each additional meal cooked with a clean stove in the last week and smoke exposure. However, we do not observe a relationship between meals and health for either the full sample of children or those who are aged five and under. The signs of the coefficients suggest improved health outcomes, but they would not be judged to be statistically significant by conventional criteria, despite the relatively large sample sizes.

Overall, Table 3 suggests that women who use cleaner stoves are in better health, but we cannot conclusively state whether there is a correlation between the stoves and children's health. Despite controlling for basic demographic characteristics, these estimates may still be biased due to a number of unobservable factors. For example, those who choose to use a clean stove may generally value health more than those who do not and thus may also undertake other health investments, either of which would lead to better health. In this case, our estimated coefficients would be biased upwards. Alternatively, the improved stoves may be disproportionately used by the sick, which would cause the estimated relationships to be biased downwards.

IV. EMPIRICAL FRAMEWORK AND EXPERIMENTAL VALIDITY

A. Empirical Framework
The experimental design allows us to solve these endogeneity problems by comparing winners and losers from the stove lottery. We begin by estimating the reduced form effect of winning the

stove on a series of outcomes, including stove use, CO exposure, health, and other non-health stove outcomes (such as fuel use and cooking time). Specifically, we estimate:

$$(1) \ Y_{ihvt} = \beta_0 + \beta_1 T_{ihvt} + (\delta_v \times \gamma_t) + \varepsilon_{ihvt}$$

where Y_{ihvt} is the outcome of interest for individual i in household h in village v at time t. T_{ihvt} is an indicator variable that equals 1 if the household was in the treatment group at time t. As we stratified the sample by village during the randomization, and treatments and control households were surveyed at about the same time within each village, we include village × survey month-year fixed effects $(\delta_v \times \gamma_t)$, i.e., there are separate fixed effects for all observations from a village in a given month-year (e.g., January 2010).[18] For CO exposure, health, and non-health stove outcomes (when possible) we additionally include the baseline value of the outcome to gain additional precision.[19] β_1 is our key parameter of interest; the random assignment of T_{ihvt} ensures that β_1 will be an unbiased estimate of the effect of being offered a stove.

To fully exploit the four years of follow-up, we additionally estimate how the treatment effect varies over time. The effect of being offered a stove may change throughout time for a variety of reasons. The effect may decline over time if the stoves break or fall into disrepair, proper use declines, or if individuals feel healthier and compensate with other unhealthy behavior (i.e., smoking). Alternatively, the effect may increase if households learn how to use the stoves better, or use them more as they learn about the benefits of the stoves over time. To capture this change, we interact the treatment effect (T_{ihvt}) with a set of indicator variables (I_k) for whether the observation falls within a given year after stove distribution $k=\{1,2,3,4\}$:

$$(2) \ Y_{ihvt} = \beta_0 + \sum_{k=1}^{4}(\beta_k(T_{ihvt} \times I_k)) + (\delta_v \times \gamma_t) + \varepsilon_{ihvt}$$

In equation (2) there are now four parameters of interest $(\beta_1, \beta_2, \beta_3 \ and \ \beta_4)$, which capture the effect of having won the lottery within one year of the stove being built, within 13 to 24 months of the stove being built, etc. Due to the timing of Lottery 2 and the surveys, β_1 is identified from winners of both Lottery 1 and 2, but the other β's are only identified from the Lottery 1 winners.

---

[18] The survey month also controls for the survey wave in which the data was collected. The results are virtually unchanged when we control for survey wave x village. These results are available from the authors upon request.
[19] We have also run all of the models without the baseline variables. The results, both in terms of the magnitude of the coefficient estimates and statistical significance, are unchanged. These results, omitted for brevity, are available from the authors upon request.

Finally to scale the results, we also estimate the effect of using any type of low-polluting stove on CO exposure using an instrumental variables strategy. We estimate:

$$(3)\ Y_{ihvt} = \theta_0 + \theta_1 Use_{ihvt} + (\delta_v \times \gamma_t) + \varepsilon_{ihvt}$$

where Use_{ihvt} is a either a measure of whether the household owns a low-polluting stove or the number of meals cooked with a good condition, low-polluting stove over the last week. As selected individuals may choose whether to take up a stove, an OLS estimate of θ_1 would be biased. Thus, we use the treatment variable (T_{ihvt}) as an instrument for Use_{ihvt} in equation (3).

Finally, note the following specification details. First, for some of the subsequent analysis, we allow the treatment effects to differ between the Lottery 1 and 2 treatment groups. For example, it is possible that households updated their beliefs on the stoves based on their neighbors' experiences, and so take-up and use may have differed between the lotteries. Second, the household level equations are weighted to account for household splits and mergers. Third, for all regression analysis, the standard errors are clustered at the household level, which is the unit at which the treatment was assigned.

B. Verification of Experimental Validity

There are two primary threats to the empirical design. First, the randomization may have produced imbalanced groups either by chance or if the randomization process was somehow corrupted. It is unlikely that the process was corrupted as the lotteries were publically conducted and our research team monitored each of them. Nonetheless, in Appendix Table 2A and 2B, we provide test results of the randomization across baseline demographics, stove use, and health for the primary cooks and children across Lottery 1 winners, Lottery 2 winners, and those who lost both lotteries. The groups are well-balanced across the 59 baseline characteristics that we consider, with only 10 percent of the differences across groups significant at the 10 percent level or more (as predicted by chance). Further details are described in the appendix.

Second, poor areas are often characterized by seasonal migration. Moreover, individuals may not have been home when our enumerators visited them if they were working in the fields, etc. Attrition would be most problematic if it is correlated with treatment status (e.g., households that obtained a new stove were less likely to migrate). We tried to minimize overall attrition by revisiting households that we could not initially locate, as well as conducting the surveys in the

evening when individuals were likely to be at home. Nonetheless, results in Appendix Table 3 (details are provided in the appendix) fail to produce any meaningful evidence of differential attrition across the treatment and control groups, implying that differential attrition is not a source of bias in the subsequent regressions.

V. RESULTS

This section is broken into four segments. We begin by examining the relationship between stove ownership and usage (Section A). Next, we explore the relationship between treatment status and CO exposure (Section B) and health (Section C). Finally, we explore the relationship between the stoves and household expenditures on fuel and repairs (Section D).

A. Improved Stove Ownership and Use

In many laboratory studies, improved stoves are evaluated in ideal conditions, with the stoves in flawless condition and used properly. However, the health gains and fuel savings that are observed in the laboratory may not be achieved in the real-world if households choose not to take up a stove, the stove is not kept in good condition, the stove is not used properly, or the stove is not used regularly (see Figure 1 for the causal chain). Moreover, the initial treatment effect may change over time, as stoves may deteriorate, and/or individuals may update their beliefs about how to use the stoves or their expected benefits. Therefore, it is of vital importance to understand how households use these technologies in practice under normal field conditions.

We begin the analysis by exploring the effect of treatment status (i.e., being offered a GV improved stove) on stove ownership and use over time. Figure 2 plots the β_k's and their 95 percent confidence intervals from a specification where we interact the treatment status indicator with indicators for months since stove construction in the village in six-month intervals, after adjustment for the village $\times$ survey month-year fixed effects (this is a modified version of equation (2)). We explore the effect of treatment status on whether households currently owned an improved GV stove (Panel A), currently had any type of low-polluting stove (Panel B), and currently had an improved stove in good condition, defined as an enumerator observed measure (Panel C). The effect of the treatment status on the number of meals cooked over the last week with any type of good condition, low-polluting stove is documented in Panel D.

16

As shown by Panel A, over 70 percent of households that won Lottery 1 built a GV stove during the first six months of the program. In the first six months, Lottery 2 winners were modestly less likely to take up a stove than Lottery 1 winners, but had the same level of stove construction by 6 to 12 months after the initial construction period in their village.[20] The fraction of households with installed GV stoves, regardless of their condition, began to decline to less than 50 percent. In the final year, the rate of stove ownership increased again as Gram Vikas made an effort to repair broken stoves from Lottery 1 during the construction of stoves for Lottery 2 winners. Appendix Table 4 illustrates that the reasons for not building/rebuilding a stove changed over time as households learned about the stoves: the fraction of households who claimed they were not interested in the stove increased from 7 percent in Year 1 to about 26 percent in Year 4. Further, the fraction of households that destroyed their stove, presumably to create space in their homes, increased from 2 percent in Year 1 to 32 percent by Year 4.

The stove condition may deteriorate over time due to normal wear and tear, coupled with insufficient maintenance. This deterioration could lead to increased levels of smoke within the household. Figure 2, Panel C reveals that the percent of treatment households with a GV improved stove in good condition was about half of that expressed in Panel A. Appendix Figure 5 helps to explain this finding by showing that the percent of Lottery 1 household winners that report ever having had a crack in the stove was 74 percent and the comparable figure for Lottery 2 is 67 percent, which is striking since they were followed for only one year. Overall, at any point in time, treatment households were between 30 to 50 percentage points more likely than control households to have had a good condition GV stove.

The presence of a good condition, low-polluting stove does not guarantee that it will be used extensively. Households may prefer traditional cooking methods and the improved stoves' double-pot technology may, for example, make it more time-consuming to cook just one pot of water. Indeed, even though about a quarter of households owned a low-polluting stove in the baseline, more than 90 percent of meals reported in the baseline were still cooked on a traditional stove.

---

[20] The reasons for not taking up a stove varied, as shown by Appendix Table 4. In the first year, about 28 percent who chose not to take a stove did so because the stoves were inconvenient: either they did not believe that they had sufficient kitchen space or the fact that the stove was not the right fit for their family size. Only 6 percent claimed that they were not building it because they had a better stove. About another quarter claimed that they were planning on building a stove soon.

Consequently, we also explored how the offer of the GV improved stove changed the number of meals cooked with any type of low-polluting stove in good condition over the previous week. This outcome captures the intensity of use and is the most direct measure of the improved stoves' potential impact on health. The Figure 2, Panel D graph shows that treatment households cook about three more meals a week than the control households on a good condition, low-polluting stove during the first year.[21] Like stove ownership, the effect falls over time, and picks up in the fourth year when there was a big push by Gram Vikas to construct and retrain households during the Lottery 2 construction.

Table 4 more formally tests the effect of the stove offer on each of these four variables. Panel A provides estimates of the overall treatment effect (equation (1)) and Panel B provides estimates of the overall treatment effect by years since stove construction in the village (equation (2)). Note that because the stove use behavior of Lottery 1 and Lottery 2 winners was not significantly different, we group them together for the remainder of the analysis.

The estimates illustrate that take-up was far from universal and proper usage was substantially smaller than take-up. About 6 percent of the control group took up the GV improved stoves and the treatment group was 62 percentage points more likely to have a GV stove than the control group (Column 1, Panel A).[22] Considering all low-polluting stoves, the overall treatment effect falls to about 47 percentage points, as about a quarter of the control group had a low polluting stove of any type (Column 2, Panel A). However, as Column 3 reports, many households did not undertake the investments necessary to keep the stoves in good condition: the treatment effect on the proportion of GV improved stoves in good condition is 36 percentage points over the entire period. The effect again is high in the first two years, falls in Year 3, and increases again during the big push in Year 4 (Column 3, Panel B). Ultimately, households did not use the stoves regularly. On average, treatment households cooked about three more meals per week (or about 20 percent more) on any good condition, improved stoves out of a total of 14 cooked meals per week.

---

[21] Looking at improved stoves in Ghana, Burwen and Levine (2011) also find that individuals do not completely reduce their use of the traditional stoves when given an improved stove, with the treatment group using an average of 1.4 traditional stoves as compared to 1.9 in the control group. In fact, they returned to three of eight villages about eight months after the stove installation and found that only about half of the improved stoves remained in regular use (i.e., warm to touch or contained reasonable amounts of ash).
[22] The overall take-up rate is not inconsistent with other preventive health products that have demonstrated health effects (see Dupas 2011 for a discussion).

If households do not use the stoves correctly, each additional meal cooked will not reduce smoke inhalation to the fullest possible extent. For example, a failure to cover the second pot opening when it is not in use will allow smoke to enter the kitchen through this opening. Similarly if households fail to clean the chimney regularly, it will become blocked and smoke will enter the kitchen when the improved stove is used.[23] It is difficult to measure proper use. Often, use is gauged through controlled kitchen tests, but households may use the stove correctly when they are being observed by researchers even if they do not typically use it properly. Similarly, self-reported measures of use may be biased upwards if households feel judged by the enumerators.[24]

Nonetheless, we collected self-reported measures of proper use. As Figure 3 shows, for the sample of those who own a stove in good condition, only about 60 percent report that they use the stoves properly, where proper use is defined as cleaning the stove in the last week, using the stove in the last week, not elevating the cook pot during use, and using the two pots correctly.

In summary, we find that stove behavior and use in real-world settings differs considerably from controlled laboratory tests. Take-up of GV stoves was only about 60 percentage points higher in treatment households than control ones, despite the fact that the stoves were highly subsidized. The share of households that maintained an improved stove in good condition was substantially smaller at 36 percentage points, and out of these, 40 percent self-reported that they did not properly clean and use the stoves to minimize indoor air pollution. In practice, treatment households continued to also use their traditional stoves, cooking only about three extra meals per week on any type of low-polluting stove in good condition.

B. Effects on Smoke Inhalation

We now test whether being offered a stove caused changes in smoke inhalation. Following Diaz, Bruce, Pope, Lie, Diaz, Arana, Smith, and Smith-Sivertson (2007), we measured Carbon Monoxide (CO) in exhaled breath to measure smoke inhalation. As discussed in the data section,

---

[23] Note that self-reported good use does not necessarily mean that the stove will be in good condition: Dutta, Shields, Edwards, and Smith (2007) find that even when households self-reported regular cleaning by dropping sand bags from the top of the chimney, the chimneys could often became clogged four to five months after installation if the cleaning was not done properly.

[24] Another reason that smoke inhalation may not be reduced to the fullest extent possible is if the stoves induce individuals to cook inside and the smoke exposure from a clean stove inside is worse than the smoke exposure from a traditional stove outside. However, we find no evidence that treatment households increased the number of meals cooked indoors.

CO is a biomarker of recent exposure to air pollution from biomass combustion, and therefore it can be used to proxy an individual's personal exposure to smoke from cooking stoves.

Table 5 and Figure 4 provide a reduced form analysis of the effect of stoves on smoke exposure for women, for those who identified themselves as primary cooks in the baseline, and for children who were old enough to be tested. Note that all specifications include the baseline values of the outcome variable and village × survey month-year fixed effects, and that the standard errors are clustered at the household level.[25]

On average, we observe limited effects on the CO concentrations in respondents' breath. For women, CO levels fall slightly, but the effect is not significant (Table 5, Panel A, Column 1). The magnitude of this effect is 1.5 percent of the mean and 1.6 percent of a standard deviation. As Pitt, Rosenzweig, and Hassan (2010) discuss, indoor air pollution is unlikely to be evenly distributed within the household, with the highest incidence likely borne by those who do most of the cooking. Therefore, we estimate the effect for the primary cook in the household at the time of the baseline. The point estimate for the average effect for primary cooks is about double that for females, but still not statistically significant over the entire period (Table 5, Panel B, Column 2). In terms of magnitude, it is small, as it is 3.1 percent of the mean and 3.6 percent of a standard deviation.[26]

These estimates mask considerable heterogeneity over time. We observe a meaningful reduction in primary cook's CO breath concentrations during the first year. Specifically, we find a 0.52 ppm reduction (7.5 percent of the control group's mean) in the CO concentration during Year 1 for primary cooks relative to the control group, when stove usage is at its highest. Thus, to the extent that they were used in the first year, they were effective in reducing CO, which supports ARTI's laboratory results that these stoves can be effective at reducing exposure to indoor air pollution. However, as usage declined so did the effect on CO: the treatment effect

---

[25] If the baseline value is missing, we assign the average of the baseline variable. We additionally include an indicator variable that equals one when the baseline value for an individual was imputed.

[26] In principle, the effect on CO concentrations of exhaled breath could be mitigated by two forms of spillovers from the treatment to the control group. First, treatment households could conduct all the cooking for the control group since they own the improved stove. The data are inconsistent with this possibility as the total number of meals cooked by treatment and control households was not significantly different during the experiment (the magnitude of the coefficient estimate is near zero, and in fact, negative). Moreover, the number of people whom the treatment household cooked for was not significantly different than that of the control households during these meals. Second, the experiment may cause control households to learn about the dangers of indoor air pollution, which leads them to change their cooking habits to protect themselves from smoke. Using data from our midline survey, we find no difference in the minutes spent cooking at arm's length from one's cooking stove, suggesting that control households were not differentially trying to protect themselves from the smoke.

for primary cooks falls to -0.17 ppm by the second year of stove ownership and is no longer significant. While smoke exposure generally decreased for women (Column 1) and children (Column 3) in the first two years, this effect is not statistically significant at conventional levels.[27]

To interpret these results, the odd columns in Table 6 report the results from estimating the effect of owning any type of low-polluting stove on CO exposure with the instrumental variables approach outlined in equation (3). Additionally, the even-numbered columns estimate the effect of an additional meal cooked on a good condition, low-polluting stove on CO exposure with the same instrumental variables approach. Columns 1 and 2 estimate the effect for all women, Columns 3 and 4 estimate the effect for the primary cooks, and Columns 5 and 6 estimate the effect for children. Note that all specifications include baseline values of the outcome variable and village × survey month-year fixed effects and are clustered at the household level.

Before turning to the results, it is worth noting that the instrumental variable estimates are not equal to the ratios of the relevant reduced form relationships in Tables 4 and 5. This is because the household-level data on the presence of a stove and meals cooked with a low-polluting stove in good condition were collected in a different survey than the individual-level data on the CO breath concentrations. Consequently, the samples differ between Table 5 and 6.

On average, owning at least one of any low-polluting stoves reduces CO levels by -0.286 ppm for women, -0.564 ppm for primary cooks, and -0.478 ppm for children. These scaled estimates suggest declines of 4.1 percent, 7.9 percent, and 8.8 percent, respectively, in smoke exposure from owning an improved stove, but none of them are statistically different from zero (Table 6, Panel A). Owning an improved cooking stove in the first year reduces CO exposure for primary cooks by -0.898 ppm, or 12.5 percent, relative to the control group (Table 6, Panel B, Column 3). However, by Year 2, this falls to -0.468 ppm and is no longer statistically significant. By Years 3 and 4, the effect becomes positive and remains statistically indistinguishable from zero.

---

[27] We also explored whether the stoves were more effective during monsoon season when individuals are more likely to cook inside. There is evidence that CO concentrations for children and primary cooks are lower in treatment households (compared to controls) during the monsoon, however these effects would not be statistically significant at conventional levels.

A comparison of these IV results with the RESPIRE studies' estimates helps to underscore the fundamental differences in approach and meaning of the studies' results. With weekly maintenance and instruction on proper use, as well as the use of stoves for most meals, the RESPIRE intervention produced a reduction in CO exposure of about 60 percent for women and 50 percent for children. This effect is much larger than the statistically insignificant 8.8 percent reduction in CO concentrations for children (the group with the largest reduction in our study) that arise from stove ownership within our study. The fact that the stoves were not used for all meals in our setting may be responsible for the differences in CO. If households had cooked all meals with an improved stove, there would have been an estimated 19.1 percent (14 meals $\times$ -0.074) reduction in CO concentrations for children. As we emphasized above, while we cannot be certain that the laboratory effect of our study's stoves are exactly equal to the effect of the RESPIRE study's stoves, it seems safe to conclude that the deterioration of the stoves over time, coupled with improper use (e.g., not covering the second pot), may be responsible for the differences in observed levels of smoke exposure reduction.

An alternative explanation for the small estimates in Table 6 is that the GV stoves cause individuals to feel healthier, which leads them to choose activities (like increased cigarette smoking) that would expose them to smoke from other sources. In this case, the impacts on smoke inhalation would be partially or even completely undone by individuals' compensatory responses. However, very few women (0.2 percent) and primary cooks (0.7 percent) report smoking during the course of the study, and therefore, changes to such rates appear unlikely to affect the overall CO results. It is also possible that men may smoke more in the household, inducing higher rates of secondhand smoke to women. Appendix Table 5 shows the reduced form effect of the treatment on the male propensity to smoke and finds no overall difference (Panel A) and no change over time (Panel B). Thus, the stoves do not appear to induce compensatory behavior that undoes the stoves' beneficial impacts.

C. Health Outcomes

This subsection reports the impact of the treatment on a wide range of health outcomes. The results in the previous subsection suggest that sustainable health effects are unlikely to operate through the channel of reduced smoke inhalation, as there are no sustainable effects on measured smoke inhalation over time. Nevertheless, it is possible that there are unobserved household

compensatory responses to the stoves that loosen budget constraints in a way that directly improves health.

For the sample of primary cooks, Figure 5A plots the effect of treatment status over time for several key health outcomes. Specifically, we examine the effect on respiratory health, as measured by the spirometer (FEV1 and FEV1/FVC × 100), as well as self-reported measures ("cough or cold" and "any illness"). The estimates are taken from the fitting of a specification that includes baseline values of the outcome variable and village × survey month-year fixed effects. The standard errors are clustered at the household level.

Larger spirometry readings indicate greater lung functioning. Even in the early period when there is a reduction in CO breath concentrations for primary cooks, no impact on lung functioning is observed. All four of these figures are remarkable in that the response function hovers around the zero line throughout the four years of follow-up: it is evident that there is little basis to reject the null hypothesis of no effect.

Table 7A reports on the formal regression analysis of the effect on health for a wider range of health outcomes for primary cooks (Panel A), children (Panel B), children aged five and under (Panel C), and infants (Panel D). As shown in the graphs, overall, we find no effect of randomly being offered a stove on health outcomes. Out of 44 health estimates, five (11 percent) are significant at the 10 percent level, which is what would be expected by chance. All five of the statistically significant effects have a counterintuitive signs, suggesting that the stove offer causes worse health, further underscoring that treatment status appears unrelated to health.[28]

In the presence of so many outcome variables, it can be informative to summarize the results by estimating an average treatment effect across the multiple outcomes. To do this, we standardized all of the outcome variables to have a mean of zero and standard deviation of 1, took the average across all outcomes for each observation, and then estimated the effect of treatment status.[29] The results are presented in Table 7B for primary cooks (Column 1), children (Column 2), and children aged five and under (Column 3). Not only are none of the estimated effects significant, they are practically very small (Table 7B, Panel A). For example, the

---

[28] We also estimate the health effects over time for each of the variables in Table 7A (omitted for brevity). This may be of especial interest for the primary cooks because of their decline in CO concentrations in the first year. However, overall, there is little evidence of health effects either in the first year or subsequent years.

[29] An observation may comprise a different number of variables due to missing data or due to the fact that surveys may have been conducted at different times. Therefore, we weight each observation by the number of variables that contribute to the average. The results from unweighted regressions are qualitatively the same.

treatment results are a -0.01 standard deviation change in health across all variables for primary cooks. Furthermore, none of the effects significantly change over time (Table 7B, Panel B).

D. Monetary and Time Costs of Improved Stoves and Self-Reported Satisfaction

Table 8 examines whether the treatment status causes changes in the monetary and time costs of using and maintaining a household's stoves. Improved stoves can affect expenditures on a number of dimensions. First, when properly used in controlled conditions, the ARTI stoves require less wood and households received training from Gram Vikas on how to achieve these fuel reductions. As such, the stoves may reduce energy use and hence fuel costs. Second, if the stoves are more efficient both in terms of heating up quickly (e.g., time required to boil water) and the two-pot functionality, cooking times may be reduced. Finally, the new stoves may alter the time spent making repairs. As recognized by the Global Alliance for Clean Cookstoves, these factors are important for adoption (particularly if households are asked to pay for the stoves). Moreover, if the stoves reduce energy use then carbon credits could be used to finance them, which is one of the avenues that is currently being explored to make them more widely available.

On average, households seem to have been convinced that they should use less wood in the new stoves: more than 60 percent of households report that they believe that GV stoves use less wood (Appendix Figure 6). However, looking at *actual* use in Table 8, wood use appears unchanged (Table 8, Column 1), while total fuel expenditures increases, although the increase is only statistically significant in Year 4 (Table 8, Column 2). The discrepancy between the laboratory test and the actual expenditures by the households may be due to improper use, or the fact that households now use both the traditional and the improved stove, perhaps simultaneously. Burwen and Levine (2011) observe a similar effect for the type of stove that they evaluate in Ghana. After eight weeks, households took less time and fuel to cook a meal in a carefully controlled test, but there was no significant decline in the actual fuel used by the family.[30] These results underscore that using laboratory or engineering tests to justify fuel efficiency gains for carbon credit calculations has the potential to be extremely misleading.

---

[30] The discrepancy between self-reports and actual outcomes has been observed in other contexts as well, and probably reflect social desirability bias, as households do not want to be impolite to people they perceive to be associated with the program (for a discussion, see Kremer, Leino, Miguel, and Zwane, 2011). In the stove context, Boy, Bruce, Smith and Hernandez (2000) report that local women in Guatemala stated that the improved cooking

Similarly, most households believe that the stoves reduce cooking time (Appendix Figure 6). However, we find that, if anything, the stoves increased the time spent cooking evening meals by about four minutes (Table 8, Column 3), although this is not statistically significant at conventional levels.

Finally, we examine the total repairs to both the improved and traditional stoves. Control households state that they repair their stoves about once a month. Treatment households made, on average, about 2.5 more repairs to their stove in the last year (Table 8, Column 4), translating to about 4.5 hours of time over the last year (Table 8, Column 5). These two effects are economically large, implying increases of 20.2 percent and 68.7 percent, respectively, and are statistically significant.[31]

Despite the fact that GV stoves increase household costs and fail to improve health, households generally report that they are satisfied with the stoves (Table 9). On a scale from one to ten with one being the best, those who obtained an improved stove rate their satisfaction with it at 2.87, with 89 percent of households happy to recommend the stoves to others. The top reasons for recommending the stoves include that they emit less smoke in the household, the household belief that they require less fuel, the two-pot functionality, and the households belief that they require less time to cook. The gap between the satisfaction results, where responses do not affect respondents' well-being (except indirectly by trying to please the surveyors), and the consequential health and CO results underscores the limitations of self-reports generally and especially in trying to learning about how individuals value new technologies.

VI. DISCUSSION

The paper's basic findings differ from the naïve OLS estimates of the impact of improved stoves on health, as well as the conventional wisdom about their benefits in the policy world. Following households for up to four years after they received a subsidized improved stove, we find that the stoves reduce the CO concentration in breath for primary cooks in the first year, but

stove (plancha) uses less wood than open fire stoves and that this was one of the features that they liked most about the stoves, even though standard measures of the stoves fuel efficiency tests suggested that cooking on a plancha was no more efficient than cooking with an open flame, and may have even required more time to cook.

[31] We examine which baseline characteristics correlate with increased probability of repairs in the treatment group. Interestingly, while female health status had no effect on repair propensity, families where children had lower respiratory health were more likely to make repairs to their stoves.

this disappears by the following year. We do not observe any effect on health, neither self-reported nor measured, or on proxies for greenhouse gas emissions.

It is noteworthy that despite the fact that we studied a relatively inexpensive stove model, this study's measured health effects are qualitatively similar to the findings from the RESPIRE study in Guatemala, which is considered the flagship randomized experiment of a stove program.[32] Table 10 assembles the results from six different RESPIRE papers that summarize the results for outcomes that are similar to those in this paper. The RESPIRE stove resulted in reduced smoke exposure: as Panel A demonstrates, for the subsample of women who were tested for emissions exposure, personal $PM_{2.5}$ was about 60 percent lower for the treatment group relative to the control group (McCracken, Smith, Díaz, Mittleman, Schwartz, 2007) and CO also declined by roughly 60 percent for mothers in the treatment group (Smith-Sivertsen, Diaz, Pope, Lie, Diaz, McCracken, Bakke, Arana, Smith, and Bruce, 2009).

Despite these reductions in IAP, the evidence on female health is surprisingly weak. Panel B demonstrates that like our findings, the estimates of the treatment effect on lung functioning are close to zero in magnitude and are not statistically significant (Smith-Sivertsen, Diaz, Pope, Lie, Diaz, McCracken, Bakke, Arana, Smith, and Bruce, 2009). There appears to be a modest and statistically significant reduction of blood pressure (Panel C), but this was found in the presence of substantial selective attrition (54 percent response rate in the treatment group and 71 percent in the control) and only after adjusting for control variables that increase the magnitude of the point estimate (McCracken, Smith, Díaz, Mittleman, Schwartz, 2007). The birth weight results are similar in this respect (Panel D). Finally, the evidence on self-reported symptoms is also mixed (Panel F). While women in the treatment group experienced a reduction in respiratory symptoms (cough, chronic chough, phlegm cough, phlegm, wheeze, and tightness of chest) the decline was statistically significant only for wheezing. The probability of self-reported sore eyes and headaches was reduced, but there was no effect on backaches (Diaz, Smith-Sivertsen, Pope, Lie, Diaz, McCracken, Arana, Smith, Bruce, 2007).[33]

---

[32] The stove that we evaluated has some very appealing features in terms of choice of stoves to evaluate in that it is relatively inexpensive to construct, had promising laboratory results, and this family of stoves is currently used by more than 166 million households (World Bank 2011). The RESPIRE stove also has some appealing aspects, although it costs about US$100-150 (Diaz, Smith-Sivertsen, Pope, Lie, Diaz, Arana, Smith, Bruce, 2007) and the full cost of the very modest health improvements would also have to include weekly inspections and free maintenance.

[33] The effect on headaches was not present in either the six-month or the 12-month follow-up; it was only present in the 18-month follow-up (Diaz, Smith-Sivertsen, Pope, Lie, Diaz, McCracken, Arana, Smith, Bruce, 2007).

The effects of the stoves on the incidence of pneumonia and respiratory syncytia virus (RSV) among children are equally as disappointing as our findings (Panel E). There was no statistically significant difference in MD-diagnosed pneumonia, which is cited as "the primary outcome" in the study. Moreover, out of 10 outcomes tracked, only one (severe pneumonia as identified by a field worker) is significant at the 5 percent level in the unadjusted data. It is only after imputing the outcome variable for missing observations that the MD-diagnosed estimates becomes marginally significant (at the 9 percent level). Even in this adjusted data, the 95 percent confidence interval for the odds ratio (0.59–1.06) *excludes* the figure of 50 percent reduction in pneumonia that is cited by the GACC.

Overall, the similarity of our health results with those of the RESPIRE study is striking because the RESPIRE experiment was conducted under conditions that approach laboratory tests. For example, the treatment households were visited weekly to ensure that they were using the stove and to provide free repairs. Given the high cost of conducting these kinds of visits, they do not reflect the way that improved stoves are typically used and thus the health results are likely an upper bound on health benefits of the RESPIRE stove. This is very puzzling because the levels of IAP during cooking periods are extremely high and there is an extensive literature linking ambient air pollution to poor human health. There are several potential explanations for these results. For example, the response function between health and IAP may be highly nonlinear, potentially including thresholds, such that reducing indoor pollution, even by more than 60 percent, at such stunningly high levels may have limited effects. If this is the case, then it is possible that there is a new stove technology that that would reduce indoor pollution exposure sufficiently to lead to health gains, but we would need to test for what this inflection point should be. Alternatively, the overall health status of individuals in very poor countries may be so low that a reduction in indoor pollution would have to be accompanied by other changes to achieve health improvements. In this case, it is possible that stove interventions could be coupled with other health interventions, but again this would need to be tested.

In addition to RESPIRE, we have emerging evidence from several other randomized evaluations of stoves. Some of the evidence is positive: Bensch and Peters (2012) evaluate an improved cooking stove for about 227 households in Senegal, and find that one year after the stoves are distributed, the stoves are used for about 71 percent of meals in the treatment group,

and there are improvements in self-reported fuel and health outcomes.[34] Some of the evidence is less so. Evaluating a mud stove in Ghana, Burwen and Levine (2011) show positive effects on self-reported health eight weeks after the distribution, although there is no effect on smoke exposure. However, a year later, only about half of the stoves appeared to still be in use. Similarly, Levine and Beltramo (2010) examine solar stoves in Senegal, and find that six months after the distribution there were no differences in the amount of time spent cooking near a fire (and only a one percent decline in fuel use) since households continued to use the traditional stoves.

The implication of the literature as whole is that the evidence on improved cooking stoves is more mixed than the conventional wisdom would suggest and that, to date, there is no clear evidence that a large-scale distribution of stoves is likely to be a cost-effective solution to the problem of indoor air pollution. Additional research needs to be done. Our findings provide clear guidance for choosing which types of stoves to test and how to design the interventions and research. First, target households appear to have a relatively low willingness to pay for improved stoves and therefore the relevance of studies of expensive stoves may be limited. In our study, despite the fact that stoves were so highly subsidized, that they were essentially free, many households refused to install one (e.g., their homes may not have been big enough to accommodate the stove or they simply did not want one). In a different setting, Miller and Mobarak (2011) directly study willingness to pay and also find that even a small cost discourages take-up. In fact, one reason why inexpensive, mud stoves are often chosen by organizations is the belief that households are unwilling to purchase the more expensive stoves. Carbon credits could in principle help with the low willingness to pay, but only if stoves actually reduce greenhouse gas emissions in practice. As noted, our study failed to find any evidence of a fuel efficiency gain given typical field conditions.

Second, it is important to study stoves under the conditions that households actually use them. In the RESPIRE study, which is primarily concerned with the clinical impacts of stoves, trained fieldworkers inspected the stoves weekly for proper use and maintenance, and then arranged for repairs if needed (Smith, McCracken, Thompson, Edwards, Shields, Canuz and

---

[34] In this study, there are follow-ups visits with households on the stove at the one, two, and seven month marks. Interestingly, in this study, the improved stove moved households from cooking inside to cooking outside.

Bruce, 2010).[35] For most organizations, these weekly visits and repairs are infeasible. In our study, Gram Vikas (which is an award winning NGO) did the standard types of training and repair activities that a typical NGO would conduct in a large-scale stove program. Given these real-world conditions, use declined over time and, despite the fact that households did expend some time repairing the stoves, the number of broken stoves increased over time. If we had forced optimal use and repairs, we would have wrongly concluded a much bigger impact of the stoves than what would happen during a large-scale distribution of them.

Third, and relatedly, it is vital to follow households long enough to understand these behavioral issues surrounding use and repairs. Treatment effects may change considerably over time, as households learn about the value of the stoves and subsequently change their investments in use and maintenance. Further, households may also experience a general depreciation of the technology, as cracks developed, chimneys become clogged, and so forth; stopgap repairs to the stoves may not be sufficient to return the technology to its original efficiency. Additionally, health effects may be cumulative and not emerge for several years, so short-duration studies may miss key benefits. Finally, in our case, we found meaningful effects on CO for primary cooks in the first year. Had we ended the study after learning this, we may have projected this effect for several years when conducting benefit-cost calculations. However, in reality, this effect was short-lived.

VII. CONCLUSION

This study shows that relatively inexpensive stoves, used under real world conditions, had limited long-run impacts. The stoves reduced smoke exposure for the primary cook in the household in the first year of the study, but after normal use they subsequently had no discernible effect on exposure. The declining effect appears to be the result of stove breakages combined with insufficient investments in maintenance, reductions in the number of meals cooked with good stoves, and inappropriate use. We found no observable effects on health, even in the early years. While households overwhelmingly claimed that the stoves used less wood, fuel use remained unchanged, and if anything, somewhat increased. The lack of obvious benefits may explain why households were not interested in using the stoves optimally.

---

[35] Even despite the carefully controlled environment, Thompson, Bruce, Eskenazi, Diaz, Pope and Smith (2011) note that the effect on child health in the RESPIRE study may have been limited by improper or incomplete use.

More broadly, this study illustrates that it is critical to allow for household behavior when evaluating health and environmental technologies. Laboratory and laboratory-style field studies are important for understanding the best case scenario for a technology. However, all technologies must ultimately be used by humans who reveal their valuations through their usage and maintenance decisions.

REFERENCES

Alonso, P.L., et al. 1991. "The Effect of Insecticide-Treated Bednets on Mortality of Gambian Children." Lancet. 1: 1499-502.

Arceo-Gomez, Hanna, and Oliva. 2012. "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City," Mimeo.

Bailis, Robert, Majid Ezzati, and Daniel Kammen. 2005. "Mortality and Greenhouse Gas Impacts of Biomass and Petroleum Energy Futures in Africa." *Science* 308(5718): 98-103.

Beers, Mark H and Robert Berkow, eds. 1999. *The Merck Manual of Diagnosis and Therapy* (17th edition). West Point, PA: Merck & Co, Inc.

Bensch, Gunther, and Jörg Peters. 2012. "A Recipe for Success? Randomized Free Distribution of Improved Cooking Stoves in Senegal." Mimeo.

Binka, F. N., F. Indome, and T. Smith, "Impact of Spatial Distribution of Permethrin-Impregnated Bed Nets on Child Mortality in Rural Northern Ghana," American Journal of Tropical Medicine and Hygiene, 59 (1998), 80–85.

Burwen, Jason and David I. Levine. 2011. "A Rapid Assessment Randomized-controlled Trial of Improved Cookstoves in Rural Ghana." Energy and Resources Group, University of California, Berkeley Working Paper.

Boy, Erick, Nigel Bruce, Kirk R. Smith, and Ruben Hernandez. 2000. "Fuel Efficiency of an Improved Wood Burning Stove in Rural Guatemala: Implications for Health, Environment, and Development." *Energy for Sustainable Development* 4(2): 23-31.

Chassang, Sylvain, Gerard Padro i Miquel, and Erik Snowberg. 2011. "Selective Trials: A Principal-Agent Approach to Randomized Controlled Experiments". Forthcoming, *American Economic Review.*

Chay, Kenneth and Michael Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *Quarterly Journal of Economics*, 118(3): 1121-1167.

Chay, Kenneth and Michael Greenstone. 2003. "Air Quality, Infant Mortality, and the Clean Air Act of 1970." *NBER Working Paper* No. 100053.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. 2012. "The Long-Run Impacts of Air Pollution on Life Expectancy: Evidence from China's Huai River Policy." Mimeo.

Currie, Janet and Matthew Neidell. 2005. "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *The Quarterly Journal of Economics* 120(3): 1003-1030.

Diaz, Esperanza, Tone Smith-Sivertsen, Dan Pope, Rolv T. Lie, Anaite Diaz, John McCracken, Byron Arana, Kirk R. Smith, and Nigel Bruce. 2007. "Eye discomfort, headache and back pain among Mayan Guatemalan women taking part in a randomised stove intervention trial." *Journal of Epidemiology and Community Health* 61:74–79.

Duggan, Mark. 2005. "Do New Prescription Drugs Pay for Themselves? The Case of Second-Generation Antipsychotics." *Journal of Health Economics* 24(1): 1-31.

Dupas, Pascaline. 2011. "Health Behavior in Developing Countries." *Annual Review of Economics* 3: 425-449.

Dupas, Pascaline. 2012. "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment," Mimeo.

Dutta, Karabi, Kyra Naumamoff Shields, Rufus Edwards, Kirk R. Smith. 2007. "Impact of improved biomass cookstoves on indoor air quality near Pune, India." *Energy for Sustainable Development* 11(2): 19-32.

Global Alliance for Clean Cookstoves. 2010. "Goals." http://cleancookstoves.org/the-alliance/goals/

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

Kremer, Michael, Jessica Leino, Edward Miguel, and Alix Zwane. 2011. "Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions."*Quarterly Journal of Economics* 126(1): 145-205.

Levine, David I. and Theresa Beltramo. 2010. "The Effect of Solar Ovens on Fuel Use, Emissions, and Health: Results from a Randomized Controlled Trial." Hass School of Business, University of California, Berkeley Working Paper.

Kandlikar, Milind, Conor C.O. Reynolds, and Andy P. Grieshop. 2009. "A Perspective Paper on Black Carbon Mitigation as a Response to Climate Change". Released by the Copenhagen Consensus Center on August 14, 2009. Retrieved December 4, 2011. http://ftc_dev.planck.mocsystems.com/uploads/tx_templavoila/PP_Black_Carbon_Kandlikar_Reynolds_Grieshop_v.1.0.pdf

Martin II, William J., Roger I. Glass, John M. Balbus, and Francis S. Collins. 2011. "A Major Environmental Cause of Death." *Science* 344: 180-181.

McCracken, John P., Kirk R. Smith, Anaite Diaz, Murray A. Mittleman, and Joel Schwartz. 2007. "Chimney Stove Intervention to Reduce Long-term Wood Smoke Exposure Lowers Blood Pressure among Guatemalan Women." *Environmental Health Perspectives* 115(7): 996-1001.

Miller, Grant and Mushfiq Mobarak. 2011. "Understanding Low Demand: An Experimental Study of Cook-stove Adoption in Bangladesh." Yale School of Management Working Paper.

Mueller, Valerie, Alexander Pfaff, John Peabody, Yaping Liu, and Kirk R. Smith. 2011. Demonstrating bias and improved inference for stoves' health benefits. *International Journal of Epidemiology 40: 1643-1651.*

Nevill, C. G., Some, E. S., Mung'ala, V. O., Muterni, W., New, L., Marsh, K., Lengeler, C. and Snow, R. W. (1996), Insecticide-treated bednets reduce mortality and severe morbidity from malaria among children on the Kenyan coast. Tropical Medicine & International Health, 1: 139–146

Phillips-Howard, Penelope A., et al. 2003. "Efficacy of Permethrin-Treated Bed Nets in the Prevention of Mortality in Young Children in an Area of High Perennial Malaria Transmission in Western Kenya." American Journal of Tropical Medicine and Hygiene. 68(Suppl. 4): 23-29.

Registrar General & Census Commisioner, India. 2001. "Census of India." http://www.censusindia.net/

Smith, Kirk R. 1987. *Biofuels, Air pollution, and Health: A Global Review*. New York: Plenum Press.

Smith. Kirk R. 2000. "National Burden of Disease in India from Indoor Air Pollution." *Proceedings of the National Academy of Sciences* 97(24): 13286–13293.

Smith, Kirk R., John P. McCracken, Lisa Thompson, Rufus Edwards, Kyra N. Shields, Eduardo Canuz, and Nigel Bruce. 2010. "Personal Child and Mother Carbon Monoxide Exposures and Kitchen Levels: Methods and Results from a Randomized Trial of Woodfired Chimney Cookstoves in Guatemala (RESPIRE)." *Journal of Exposure Science and Environmental Epidemiology* 20: 406–416

Smith, Kirk R., John P McCracken, Martin W Weber, Alan Hubbard, Alisa Jenny, Lisa M Thompson, John Balmes, Anaité Diaz, Byron Arana, and Nigel Bruce. 2011. "Effect of reduction in household air pollution on childhood pneumonia in Guatemala (RESPIRE): a randomised controlled trial." *The Lancet* 378: 1717-26.

Smith-Sivertsen, Tone, Esperanza Diaz, Dan Pope, Rolv T. Lie, Anaite Diaz, John P. McCracken, Per Bakke, Byron Arana, Kirk R. Smith, and Nigel Bruce. 2009. "Effect of Reducing Indoor Air Pollution on Women's Respiratory Symptoms and Lung Function: RESPIRE Guatemala Randomized Trial." *American Journal of Epidemiology* 170(2): 211-220.

Thomas, Duncan, Elizabeth Frankenberg, Jed Friedman, Jean-Pierre Habicht, Nicholas Ingwersen, Christopher McKelvey, Mohammed Hakimi, Jaswadi, Gretel Pelto, Bondan Sikoki, Teresa Seeman, James P. Smith, Cecep Sumantri, Wayan Suriastini, Siswanto Wilopo "Causal Effect of Health on Labor Market Outcomes: Experimental Evidence" January 2010, MIMEO.

Thompson, Lisa M., Nigel Bruce, Brenda Eskenazi, Anaite Diaz, Daniel Pope, and Kirk R. Smith. 2011. "Impact of Reduced Maternal Exposures to Wood Smoke from an Introduced Chimney Stove on Newborn Birth Weight in Rural Guatemala." *Environmental Health Perspectives* 119(10).

The World Bank. 2011. "Household Cookstoves, Environment, Health, and Climate Change: A New Look at an Old Problem." The Environment Department (Climate Change): http://climatechange.worldbank.org/sites/default/files/documents/Household%20Cookstoves-web.pdf

Table 1:   Baseline Household Demographic Characteristics and Stove Usage

	Mean	St Dev	N
	(1)	(2)	(3)
Panel A: Socio-demographic Characteristics			
Household Size	6.57	3.49	2530
Monthly Per Capita Household Expenditures	475.28	299.00	2494
Minority Household (Scheduled Caste or Tribe)	0.43	0.50	2516
Has Electricity in Household	0.47	0.50	2529
Male Head Ever Attended School	0.69	0.46	2085
Male Head Literate	0.58	0.49	2100
Female Head Ever Attended School	0.32	0.47	2369
Female Head Literate	0.20	0.40	2374
Female Has a Savings Account	0.68	0.47	2408
Panel B: Stove Ownership and Use			
Traditional Stove	0.99	0.09	2480
Any Type of "Clean Stove"	0.23	0.42	2480
Improved Stove	0.01	0.10	2480
Kerosene	0.10	0.30	2480
Biogas	0.03	0.17	2480
LPG	0.04	0.20	2480
Electric	0.11	0.31	2480
Coal	0.00	0.06	2480
Cooked Most Meals with Traditional Stove in Last Week	0.93	0.26	2453
Meals Cooked Last Week	13.75	4.08	2482
Meals Cooked Last Week with Traditional Stove	12.60	4.66	2470
Primary Cooks (% female)	0.96	0.20	2476
Panel C: Number of Meals Cooked Each Week, By Stove Location			
Open Area	7.46	7.19	2515
Semi-open Area	5.09	6.86	2515
Enclosed Area	0.86	3.18	2515
Panel D: Fuel			
Ever Used Wood as Fuel	0.99	0.12	2523
Minutes Spent Gathering Wood Yesterday (if gathered wood)	313.62	274.99	282
Wood Used for Last Meal (in kg)	4.96	6.28	1069
Meals Per Bundle of Wood	5.04	7.24	2074
Household Gathers Wood	0.83	0.37	2515
Ever Bought Wood	0.35	0.48	2468
Ever Sold Wood	0.20	0.40	2101

Notes: This table provides sample statistics on the baseline demographics characteristics and stove usage for households. The top 1 percent of values is dropped from continuous variables.

Table 2:   Baseline CO Exposure and Health

	Mean	St Dev	N	Mean	St Dev	N
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Adults		*Women*			*Primary Cooks*	
CO	7.55	6.13	3372	7.75	6.19	2017
FEV1	1.93	0.37	2755	1.97	0.37	1689
FVC	2.26	0.42	2751	2.30	0.44	1687
FEV1/FVC * 100	89.70	6.30	2685	89.62	6.11	1650
BMI	18.82	2.58	3738	18.91	2.51	2132
Cold or Cough	0.52	0.50	4253	0.52	0.50	2468
Phlegm	0.13	0.34	4243	0.13	0.34	2467
Headache	0.49	0.50	4252	0.49	0.50	2468
Sore eyes	0.30	0.46	4254	0.28	0.45	2468
Wheezing	0.02	0.13	4243	0.01	0.11	2467
Tightness in Chest	0.05	0.21	4242	0.04	0.21	2466
Any Illness	0.87	0.33	4255	0.86	0.34	2468
Health Expenditures in the Last Month	65.17	177.53	4199	61.60	169.95	2440
Panel B: Children		*All Children*			*Children Aged 5 and Under*	
CO	6.48	5.29	517			
BMI	-1.85	1.29	2700	-1.94	1.37	929
Cough	0.40	0.49	3343	0.45	0.50	1399
Consulted Health Provider About Fever	0.27	0.44	3282	0.30	0.46	1373
Earache	0.09	0.28	3343	0.09	0.29	1399
Skin Irritation	0.13	0.34	3342	0.14	0.35	1398
Vision Problems	0.01	0.11	3343	0.01	0.10	1399
Hearing Problems	0.01	0.11	3343	0.01	0.10	1399
Vomiting	0.08	0.27	3343	0.09	0.28	1399
Diarrhea	0.08	0.27	3343	0.11	0.31	1399
Abdominal Pain	0.14	0.35	3342	0.15	0.35	1398
Worms	0.09	0.29	3339	0.12	0.32	1396
Weakness	0.22	0.41	3342	0.22	0.42	1398
Any Illness	0.73	0.44	3343	0.77	0.42	1399
Health Expenditures in the last month	46.15	102.11	3249	51.62	108.99	687

Notes: This table provides sample statistics on baseline IAP and health for women, primary cooks, and children.   For continuous variables, the top 1 percent of values are dropped.   BMI for children is standardized using values from the 2000 US CDC Population of Children.

Table 3: Correlations Between the Number of Meals Cooked in the Last Week with a Clean Stove and Baseline Smoke Exposure and Health

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
							Panel A: Females								
			FEV1/FVC *100		Cough or							Health			
	CO	FEV1		BMI	Cold	Sore Eyes	Headache	Phlegm	Wheeze	Tight Chest	Any Illness	Expenditures			
Meals	-0.036	0.002	0.076**	0.038***	-0.005**	-0.004**	0.001	-0.002	-0.001***	-0.000	0.001	1.223			
	(0.034)	(0.002)	(0.036)	(0.013)	(0.002)	(0.002)	(0.002)	(0.001)	(0.000)	(0.001)	(0.002)	(1.026)			
N	3,283	2,681	2,614	3,645	4,159	4,160	4,158	4,149	4,149	4,149	4,161	4,105			
							Panel B: Primary Cooks in Baseline								
			FEV1/FVC *100		Cough or							Health			
	CO	FEV1		BMI	Cold	Sore Eyes	Headache	Phlegm	Wheeze	Tight Chest	Any Illness	Expenditures			
Meals	-0.013	0.001	0.086*	0.043**	-0.004	-0.006***	0.004	-0.001	-0.001*	-0.000	0.000	1.749			
	(0.039)	(0.003)	(0.046)	(0.017)	(0.003)	(0.002)	(0.003)	(0.002)	(0.000)	(0.001)	(0.002)	(1.320)			
N	1,967	1,654	1,617	2,092	2,421	2,421	2,421	2,420	2,420	2,420	2,421	2,393			
							Panel C: Children Aged 13 and Under in Baseline								
				Consult for					Abdominal	Hearing	Vision				Health
	CO	BMI	Cough	Fever	Earache	Skin Infection	Vomit	Weakness	Pain	Problems	Problems	Worms	Diarrhea	Any Illness	Expenditures
Meals	-0.124**	0.005	-0.002	-0.001	0.001	-0.000	-0.000	-0.002	-0.001	-0.000	0.001	0.002	-0.001	-0.004	-0.322
	(0.056)	(0.009)	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)	(0.586)
N	507	2,659	3,293	3,232	3,293	3,292	3,293	3,292	3,292	3,293	3,293	3,289	3,293	3,293	3,199
							Panel D: Children Aged 5 and Under in Baseline								
				Consult for					Abdominal	Hearing	Vision				
		BMI	Cough	Fever	Earache	Skin Infection	Vomit	Weakness	Pain	Problems	Problems	Worms	Diarrhea	Any Illness	
Meals		-0.019	-0.005	-0.003	0.001	-0.000	0.001	-0.002	-0.001	-0.000	0.002	-0.001	-0.000	-0.002	
		(0.015)	(0.005)	(0.004)	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)	(0.000)	(0.002)	(0.003)	(0.002)	(0.004)	
N		915	1,379	1,353	1,379	1,378	1,379	1,378	1,378	1,379	1,379	1,376	1,379	1,379	

Notes: This table provides the correlation between the number of meals cooked with a clean stove at time of baseline and each variable listed in the table. All regressions are estimated using OLS, include village x month of survey x year of survey fixed effects, and standard errors are clustered at the household level. In addition, all regressions include indicators for whether the household belongs to a disadvantaged minority, the household has electricity, the male household head has attended school, the male household head is literate, the female household head (or wife of male head) has attended any school and is literate, any female member has a savings account, as well as monthly household per capita consumption. For continuous variables, the top 1 percent of values are dropped. BMI for children is standardized using values from the 2000 US CDC Population of Children. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4: Reduced Form Effect of Stove Offer on Take-Up and Usage

	Gram Vikas Improved Stove at Time of Survey	Any Low-Polluting Stove	Gram Vikas Improved Stove in Good Condition	Number of Meals Cooked with any Good Condition, Low-Polluting Stove
	(1)	(2)	(3)	(4)
Panel A: Overall Treatment Effect				
Treat	0.618***	0.469***	0.364***	3.086***
	(0.011)	(0.012)	(0.011)	(0.178)
Panel B: By Months Since Stove Construction				
Treat * I(0 to 12 mo)	0.654***	0.478***	0.364***	3.493***
	(0.012)	(0.014)	(0.012)	(0.230)
Treat * I(13 to 24 mo)	0.670***	0.500***	0.430***	3.424***
	(0.014)	(0.018)	(0.017)	(0.337)
Treat * I(25 to 36 mo)	0.441***	0.396***	0.286***	1.759***
	(0.014)	(0.015)	(0.015)	(0.286)
Treat * I(37 to 48 mo)	0.722***	0.516***	0.429***	4.032***
	(0.019)	(0.018)	(0.020)	(0.325)
N	18,966	17,459	15,370	6,593
Control Group Mean	0.064	0.245	0.044	2.381

Notes: This table provides information on stove ownership and usage over time. All regressions are estimated using OLS, include village x month of survey x year of survey fixed effects, and standard errors are clustered at the household level. Regressions are weighed to account for household splits and mergers. Good condition is defined as those stoves reported to be in good condition as observed by the enumerator. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5: Reduced Form Effect of Stove Offer on CO Exposure

	Females	Primary Cooks	Children
	(1)	(2)	(3)
Panel A: Overall Treatment Effect			
Treat	-0.105	-0.223	-0.131
	(0.161)	(0.196)	(0.180)
Panel B: By Months Since Stove Construction			
Treat * I(0 to 12 mo)	-0.234	-0.524*	-0.318
	(0.235)	(0.280)	(0.288)
Treat * I(13 to 24 mo)	-0.367	-0.171	-0.107
	(0.392)	(0.490)	(0.445)
Treat * I(25 to 36 mo)	0.184	0.078	-0.161
	(0.254)	(0.317)	(0.212)
Treat * I(37 to 48 mo)	0.027	0.101	0.279
	(0.326)	(0.436)	(0.414)
N	7,293	4,232	4,400
Control Group Mean	6.901	7.129	5.456

Notes: This table provides the reduced form effect of being offered a GV stove on CO levels.   All regressions are estimated using OLS, include village x month of survey x year of survey fixed effects, include baseline CO, and standard errors are clustered at the household level.   The top 1 percent of values are dropped.   Primary cook is defined as the individual who reported, in the baseline survey, cooking the majority of meals in the household during the last week. *** p<0.01, ** p<0.05, * p<0.1

Table 6: IV Effect of Stove Usage on CO

	Females		Primary Cooks		Children	
	Any Type of Low-Polluting Stove	Meals on Good Condition, Low-Polluting Stove	Any Type of Low-Polluting Stove	Meals on Good Condition, Low-Polluting Stove	Any Type of Low-Polluting Stove	Meals on Good Condition, Low-Polluting Stove
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Overall Treatment Effect						
Stove Variable	-0.286	-0.029	-0.564	-0.064	-0.478	-0.074
	(0.367)	(0.054)	(0.421)	(0.064)	(0.483)	(0.075)
Panel B: By Months Since Stove Construction						
Stove Variable x I(0 to 12 mo)	-0.422	-0.064	-0.898*	-0.107	-0.658	-0.132
	(0.458)	(0.061)	(0.522)	(0.071)	(0.697)	(0.094)
Stove Variable x I(13 to 24 mo)	-0.593	0.006	-0.468	-0.001	-0.205	0.029
	(0.855)	(0.140)	(0.986)	(0.173)	(0.956)	(0.189)
Stove Variable x I(25 to 36 mo)	0.438	0.098	0.299	0.048	-0.213	-0.008
	(0.698)	(0.149)	(0.848)	(0.190)	(0.639)	(0.118)
Stove Variable x I(37 to 48 mo)	0.193	-0.017	0.044	-0.016	0.261	0.015
	(0.684)	(0.076)	(0.807)	(0.092)	(0.979)	(0.118)
N	7,105	6,784	4,043	3,863	4,098	3,901

Notes: This table provides the coefficient estimate of the effect of stove usage on CO levels, where stove usage is instrumented by treatment status. All regressions include village x month of survey x year of survey fixed effects, include baseline CO, and standard errors are clustered at the household level. The top 1 percent of values are dropped. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 7A: Reduced Form Effect of Program on the Health

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Panel A: Primary Cooks															
	FEV1	FEV1/FVC * 100	Cough or Cold	Sore Eyes	Headache	Phlegm	Wheeze	Tight Chest	Any Illness	BP - Systolic	BP - Diastolic	Health Expenditures			
Treat	0.003	-0.005	0.005	-0.005	-0.005	0.006	-0.001	-0.005	-0.017	0.472	0.015	1.803			
	(0.018)	(0.003)	(0.015)	(0.010)	(0.015)	(0.008)	(0.002)	(0.004)	(0.014)	(0.579)	(0.388)	(4.205)			
N	3,103	3,068	5,337	5,336	5,336	3,580	3,578	3,578	5,336	3,051	3,051	5,141			
Control Group Mean	1.922	0.859	0.410	0.111	0.358	0.054	0.005	0.018	0.753	110.988	72.986	45.888			
Panel B: Children Aged 13 and Under in the Baseline															
	BMI	Cough	Consult for Fever	Earache	Skin Infection	Vomit	Weakness	Abdominal Pain	Hearing Problems	Vision Problems	Worms	Diarrhea	Any Illness	Health Expenditures	Days of School Missed Last Week
Treat	-0.081*	0.011	0.006	-0.001	-0.002	0.007	-0.001	0.008	-0.000	0.005***	0.003	0.002	0.005	0.838	-0.006
	(0.043)	(0.011)	(0.010)	(0.004)	(0.004)	(0.005)	(0.008)	(0.008)	(0.001)	(0.002)	(0.005)	(0.008)	(0.013)	(1.795)	(0.006)
N	7,138	9,498	10,308	9,499	9,497	6,973	6,972	6,971	6,969	6,968	6,536	6,971	9,863	9,628	3,555
Control Group Mean	-1.611	0.262	0.178	0.034	0.037	0.035	0.072	0.077	0.002	0.002	0.026	0.072	0.561	25.549	0.122
Panel C: Children Aged 5 and Under in the Baseline															
	BMI	Cough	Consult for Fever	Earache	Skin Infection	Vomit	Weakness	Abdominal Pain	Hearing Problems	Vision Problems	Worms	Diarrhea	Any Illness	Health Expenditures	
Treat	-0.117*	0.018	0.003	-0.004	0.001	0.008	0.008	0.013	-0.001	0.006***	0.006	0.006	0.005	0.512	
	(0.065)	(0.013)	(0.011)	(0.005)	(0.005)	(0.006)	(0.009)	(0.010)	(0.001)	(0.002)	(0.007)	(0.009)	(0.015)	(2.033)	
N	2,724	6,034	6,395	6,036	6,037	4,370	4,369	4,370	4,369	4,369	3,941	4,370	6,277	6,207	
Control Group Mean	-1.874	0.229	0.165	0.033	0.029	0.031	0.065	0.082	0.003	0.001	0.021	0.061	0.533	24.679	
Panel D: Pregnancy Infant Outcomes															
	Birthweight	Infant Mortality	Infant Mortality, Stillbirths and Miscarriages												
Treat	53.375	0.021**	0.018												
	(119.120)	(0.010)	(0.017)												
N	520	1,109	1,176												
Control Group Mean	2920.942	0.068	0.119												

Notes: This table provides the reduced form effect of being offered a GV stove on health. All regressions in Panel A - C are estimated using OLS, include village x month of survey x year of survey fixed effects, and standard errors are clustered at the household level. In Panel D, the mortality regressions in columns 2 and 3 include village x survey quarter x survey year fixed effects, and the birthweight regression includes village x birth quarter x birth year fixed effects. For all variables except blood pressure and days of school missed last week, we additionally include the baseline value. For continuous variables, the top 1 percent of values are dropped. BMI for children is standardized using values from the 2000 US CDC Population of Children. *** p<0.01, ** p<0.05, * p<0.1

Table 7B: The Reduced Form Effect of Stoves on Health

	Primary Cooks	Children	Children Aged Five and Under
	(1)	(2)	(3)
Panel A: Overall Treatment Effect			
Treat	-0.010	0.006	0.014
	(0.014)	(0.011)	(0.012)
Panel B: By Months Since Stove Construction			
Stove Variable x I(0 to 12 mo)	-0.010	0.005	0.011
	(0.019)	(0.014)	(0.016)
Stove Variable x I(13 to 24 mo)	0.012	0.041	0.042
	(0.029)	(0.026)	(0.029)
Stove Variable x I(25 to 36 mo)	-0.020	-0.014	0.002
	(0.023)	(0.019)	(0.022)
Stove Variable x I(37 to 48 mo)	-0.011	0.014	0.017
	(0.024)	(0.019)	(0.022)

Notes: This table provides the reduced form effect of the stove offer on the standardized indices of health.

Table 8: Time and Cost of Operating Stoves

	Total Wood Used at Last Meal (Kg) (1)	Total Fuel Costs Last 30 days (Rupees) (2)	Time Spent Cooking Last Evening Meal (Minutes) (3)	Number of Repairs Made in the Last Year (4)	Time Spent On Repairs in the Last Year (Minutes) (5)
Panel A: Overall Treatment Effect					
Treat	0.048	10.041	3.952	2.431***	277.981***
	(0.125)	(7.065)	(3.304)	(0.506)	(14.665)
Panel B: By Months Since Stove Construction					
Treat x I(0 to 12 mo)	-0.024	14.633	1.495	6.438***	60.977**
	(0.182)	(11.446)	(5.666)	(1.258)	(26.503)
Treat x I(13 to 24 mo)	0.069	-22.429	12.511*	2.140***	24.699
	(0.184)	(16.260)	(6.613)	(0.632)	(28.468)
Treat x I(25 to 36 mo)	0.245	13.332	1.226	-0.240	-23.200
	(0.248)	(11.756)	(4.514)	(0.362)	(18.782)
Treat x I(37 to 48 mo)	-0.157	22.314*	5.666	2.576**	79.193**
	(0.281)	(12.072)	(6.066)	(1.122)	(32.205)
N	5,619	4,599	4,651	3,786	3,794
Control Group Mean	3.371	327.579	163.923	12.053	404.624

Notes: This table provides the reduced form effect of being offered a GV stove on stove expenditures. All regressions are estimated using OLS, include village x month of survey x year of survey fixed effects, and standard errors are clustered at the household level. The specifications in Columns 1-3 also include baseline values. The top 1 percent of values are dropped for continuous variables. Wood use is in kilograms, time variables are in minutes, and fuel costs are in rupees. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 9: Self-Reported Satisfaction with GV Stoves

	Mean	St Dev	N
	(1)	(2)	(3)
Satisfaction with improved stove	2.87	2.25	1200
Would recommend improved stove	0.89	0.31	1192

Panel A: Reasons Why Would Recommend

	Any Reasons		
Requires less time	0.34	0.47	1085
Requires less fuel	0.58	0.49	1085
Food tastes better	0.06	0.24	1085
Less smoke	0.63	0.48	1085
Like the two-pot functionality	0.49	0.50	1085
Easier to clean	0.09	0.28	1085
Better for health of self	0.05	0.22	1085
Better for health of children	0.01	0.12	1085
Pot does not turn black	0.01	0.10	1085

Panel B: Reasons Why Would Not Recommend

	Any Reasons		
Requires more time	0.46	0.50	140
Requires more fuel	0.45	0.50	140
Food tastes worse	0.03	0.16	140
More difficult to clean	0.08	0.28	140
Have to repair it	0.05	0.22	140
More smoke than traditional	0.01	0.09	140

Notes: This table provides sample statistics on self-reported satisfaction with the GV improved cooking stoves for those who own one. The satisfaction variable is out of 10, with 1 being the highest level of satisfaction.

Table 10: Summary of RESPIRE Findings

Outcome (1)	Study (2)	Point Estimate (3)	P-Value (4)	95% Confidence Interval (5)	Point Estimate (6)	P-Value (7)	95% Confidence Interval (8)	Point Estimate (9)	P-Value (10)	95% Confidence Interval (11)
		Estimate			Adjusted Estimate			Imputed Data		
A. Smoke Exposure										
CO Passive Diffusion Tubes: Child	(1)	-52%		[- 56, -47]						
CO Passive Diffusion Tubes: Mother	(1)	-61%		[-65, -57]						
Continuous CO Monitors	(1)	-90%		[-92, -87]						
PM 2.5	(2)	-61%								
B. Lung Functioning										
FEV (6, 12 and 18 months) (5)	(5)	-0.02		[-0.09, 0.04]						
FVC (6, 12 and 18 months) (5)	(5)	-0.04		[-0.01, 0.03]						
(FEV1:FVC) *100 (6, 12, and 18 months) (5)	(5)	0.41		[-0.44, 1.27]						
C. Blood Pressure										
SBP Estimate (2)	(2)	-2.3	0.30	[–6.6, 2.0]	–3.7	0.10	[–8.1, 0.6]			
DBP Estimate (2)	(2)	-2.2	0.09	[–4.7, 0.3]	–3.0	0.02	[–5.7, –0.4]			
D. Infant Outcomes										
Mean Birth Weight (3)	(3)	68	0.28	[–56, 191]	89	0.13	[–27, 204]			
Low Birth Weight Odds Ratio (3)	(3)				0.74		[0.33, 1.66]			
E. Pneumonia and RSV Incidence for Children										
Field Worker Assessed Pneumonia Rate Ratio	(4)	0.91	0.39	[0.74, 1.13]						
Field Worker Assessed Severe Pneumonia Rate Ratio	(4)	0.56	0.04	[0.32, 0.97]						
Clinical Pneumonia Rate Ratio All	(4)	0·84	0.257	[0.63, 1.13]				0.78	0.095	[0.59, 1.06]
Clinical Pneumonia Rate Ratio hypoxemic	(4)	0.74	0.128	[0.50, 1.09]				0.67	0.042	[0.45, 0.98]
Clinical Pneumonia Rate Ratio CXR confirmed	(4)	0.87	0.586	[0.52, 1.45]				0.74	0.231	[0.42, 1.15]
Clinical Pneumonia Rate Ratio CXR hypoxemic	(4)	0.8	0.505	[0.41, 1.56]				0.68	0.234	[0.36, 1.33]
RSV(-)	(4)	0.91	0.598	[0.63, 1.30]				0.79	0.192	[0.53, 1.07]
RSV(-) hypoxemic	(4)	0.61	0.066	[0.35 ,1.03]				0.54	0.026	[0.31, 0.91]
RSV(+)	(4)	0.94	0.801	[0.59, 1.49]				0.76	0.275	[0.42, 1.16]
RSV(+) hypoxemic	(4)	1.05	0.867	[0.60, 1.83]				0.87	0.633	[0.46, 1.51]
F. Self Reported Symptoms										
Cough (4) (7)	(5)		NS							
Chronic Cough (4) (7)	(5)		NS							
Phlegm (4) (7)	(5)		NS							
Chronic Phlegm (4) (7)	(5)		NS							
Wheeze (Relative Risk) (4)	(5)	0.42		[.25, .70]						
Tightness in Chest (4) (7)	(5)		NS							
Number of Symptoms (Odds Ratio)	(5)	0.7	0.03	[.50, .97]						
% Sore Eyes in Past Month (6 Month) (6) (7)	(6)	-19.00	S							
% Sore Eyes in Past Month (12 Month) (6) (7)	(6)	-26.10	S							
% Sore Eyes in Past Month (18 Month) (6) (7)	(6)	-26.20	S							
% Headache in Past Month (6 Month) (6) (7)	(6)	0.00	NS							
% Headache in Past Month (12 Month) (6) (7)	(6)	-7.10	NS							
% Headache in Past Month (18 Month) (6) (7)	(6)	-20.30	S							
% Back pain in Past Month (6 Month) (6) (7)	(6)	-0.20	NS							
% Back pain in Past Month (12 Month) (6) (7)	(6)	-2.00	NS							
% Back pain in Past Month (18 Month) (6) (7)	(6)	-7.30	NS							

Notes:

(1) Adjusted implied controls for the number of minutes the tubes were worn.

(2) Adjusted for age, BMI, daily average apparent temperature, rainy season, day of week, time of day, use of a temascal, having household electricity, an asset index, ever smoking, SHS exposure, and a random effect

(3) Adjusted for maternal height, gravidity, maternal diastolic blood pressure, and season of birth.

(4) Information on point estimate and p-values are unavailable.

(5) Paper also reports results for just 12 and 18 months of follow-up and finds similar results. These are omitted from the table for brevity.

(6) The Mann-Whitney U test was used for testing the significance of differences.
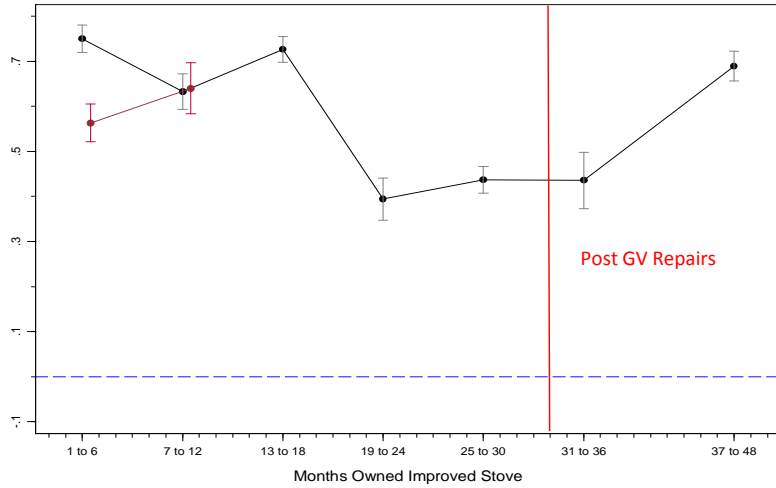
(7) "NS" means not significant; "S" means significant.
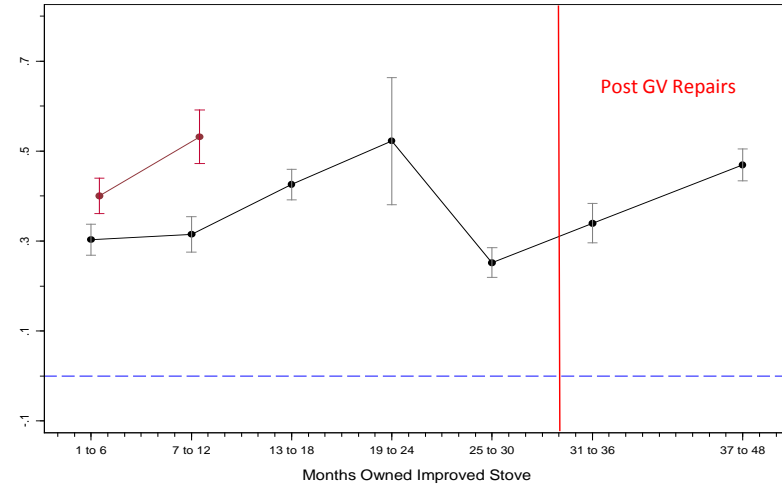
Figure 1: Causal Chain
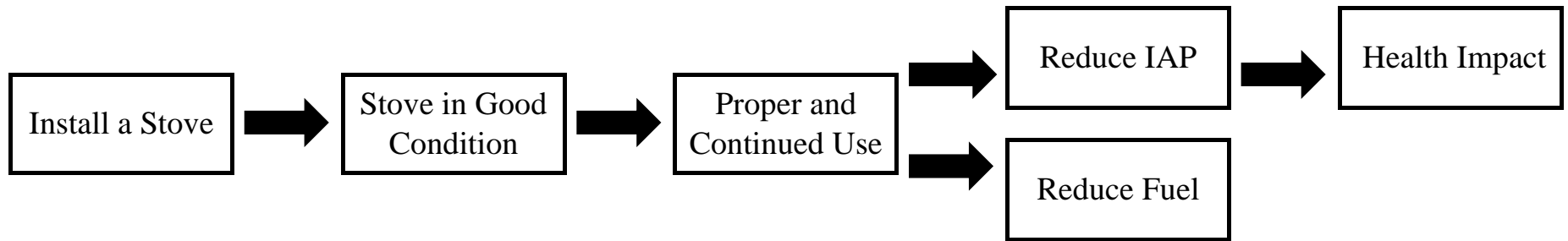
Note: This figure traces out the behavioral chain necessary to observe health and fuel impacts after a stove offer is made.
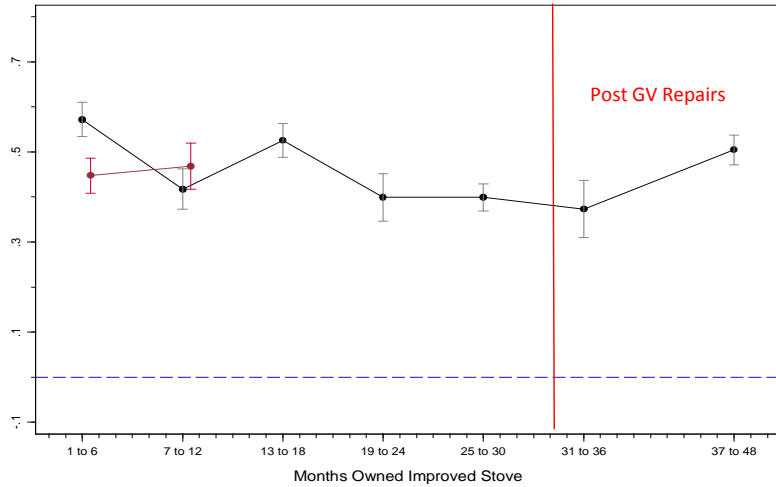
Figure 2: Stove Ownership and Usage, by Time

Panel A: Improved Stove Existing at Time of Survey
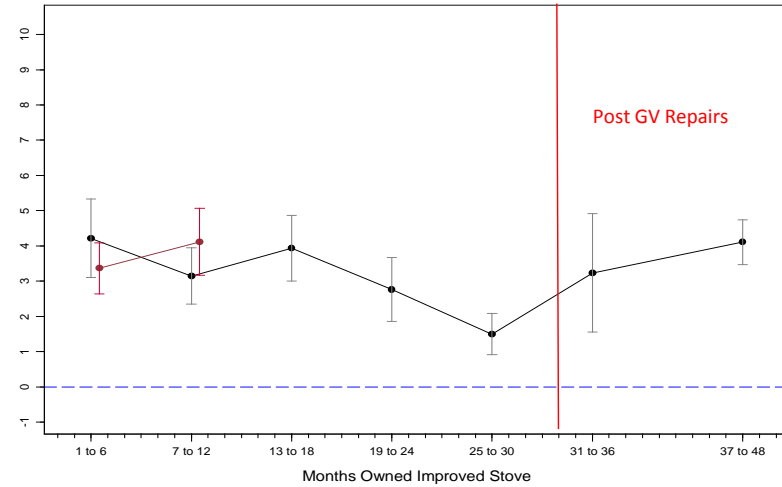

Panel C: Improved Stove in Good Condition

Panel B: Owns Any Low-Polluting Stove

Panel D: Meals Cooked on a Good Condition, Low-Polluting Stove Last Week

Notes: These figures show the difference in stove usage between the treatment and control groups, by months since stove construction in the village and lottery status, conditional on village x month of survey x year fixed effects. Regressions are weighed to account for splits and mergers. The black line signifies Lottery 1, while the red line signifies Lottery 2. The bars represent the 95th percent confidence interval.

Figure 3: Proper Use for Those Who Owned a Good Condition GV Stove
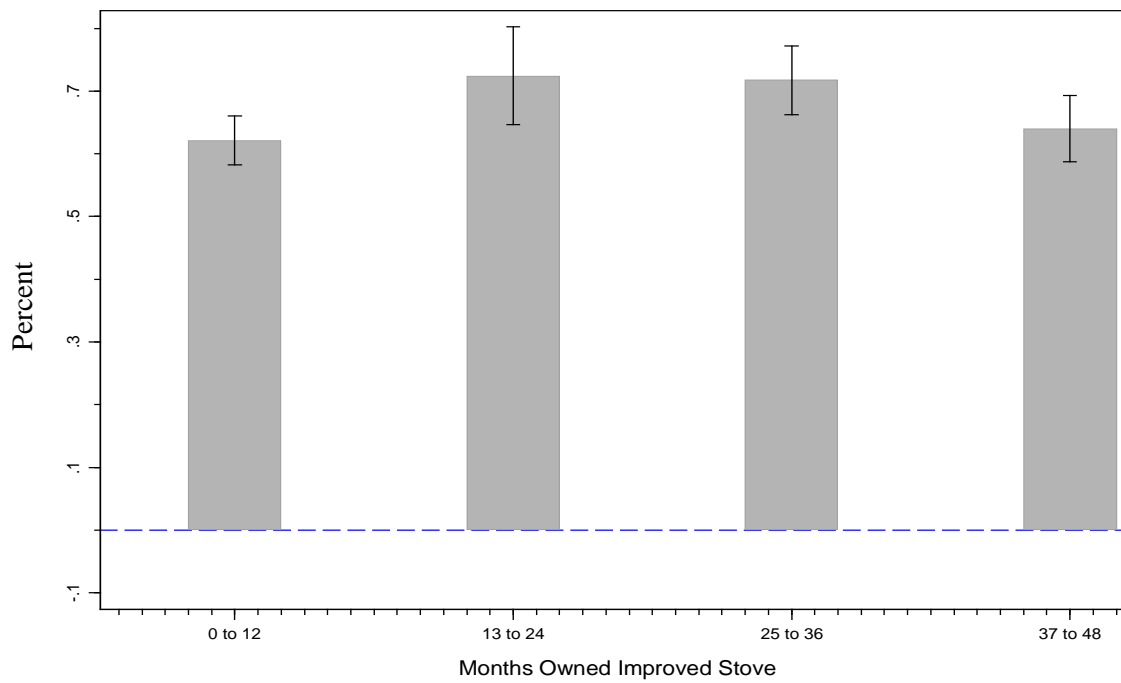
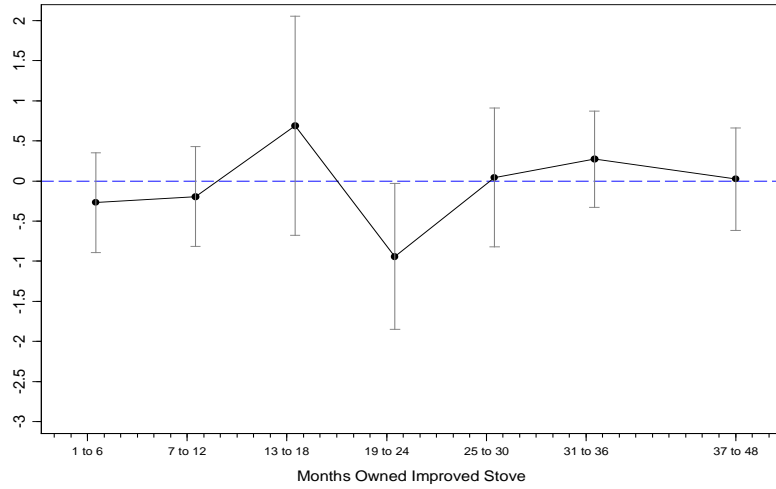Notes: Good condition is defined as observed by the enumerator. Proper use is defined as cleaning the stove in the last week, using the stove in the last week, not elevating the cookpot during use, and using the two pots correctly.

Figure 4: Reduced Form Effect on CO, by Time

Panel A: Women

Panel C: Children Aged 13 and Under in Baseline
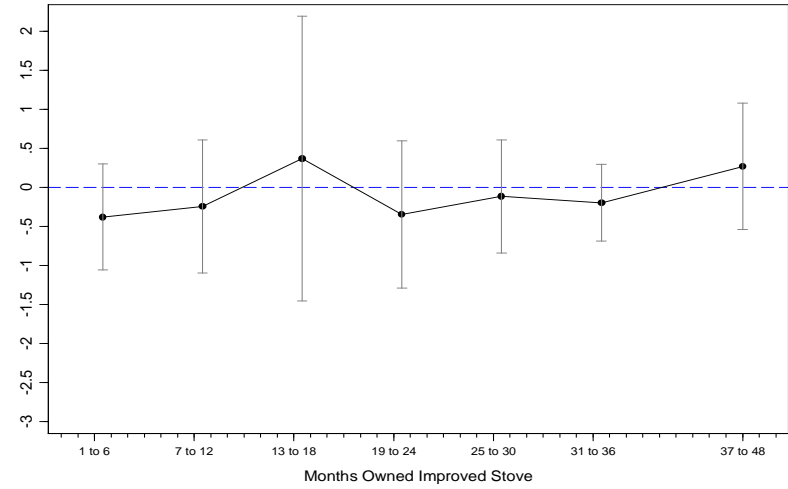
Panel B: Primary Cooks
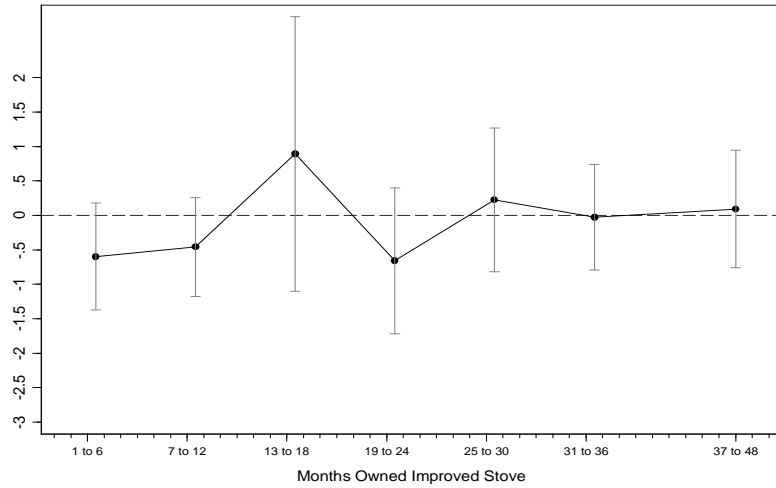
Notes: These figures show the difference in CO between the treatment and control groups, by months since stove construction in the village, conditional on village x month of survey x year fixed effects. The bars represent the 95th percent confidence interval. Lottery 1 and Lottery 2 are grouped together as treatment. The top 1 percent of values for CO are dropped. Primary cook is defined as the individual who reported, in the baseline survey, cooking the majority of meals in the household during the last week.

Figure 5A: Reduced Form Effect of the Stove Offer on the Health of the Primary Cooks, by Time

Notes: These figures show the difference in health outcomes between the treatment and control groups, by months since stove construction in the village, conditional on village x month of survey x year fixed effects. Lottery 1 and Lottery 2 are grouped together as treatment. The bars represent the 95th percent confidence interval. For continuous variables, the top 1 percent of values are dropped.

Figure 5B: Reduced Form Effect of Stove Offer on the Health Outcomes of Children Aged 13 and Under in the Baseline, by Time

Panel A: BMI
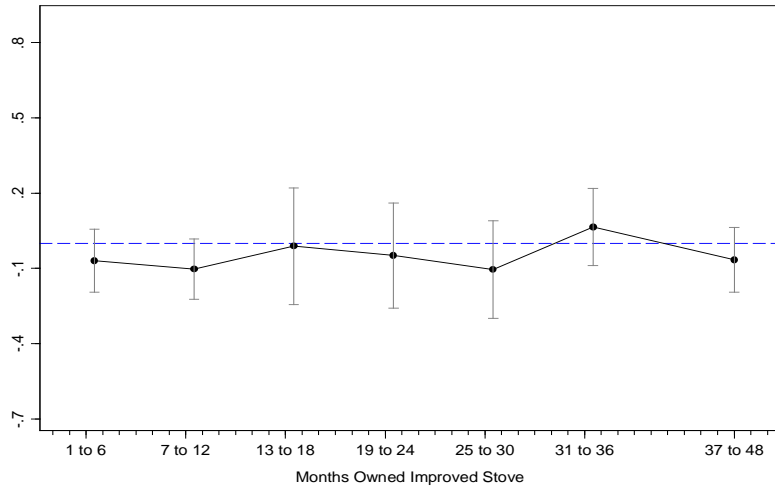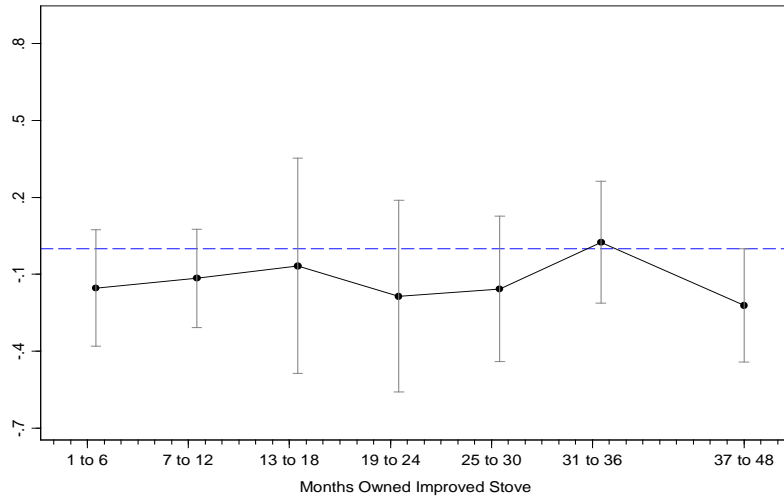
Panel C: Consulted Health Care Professional for Fever
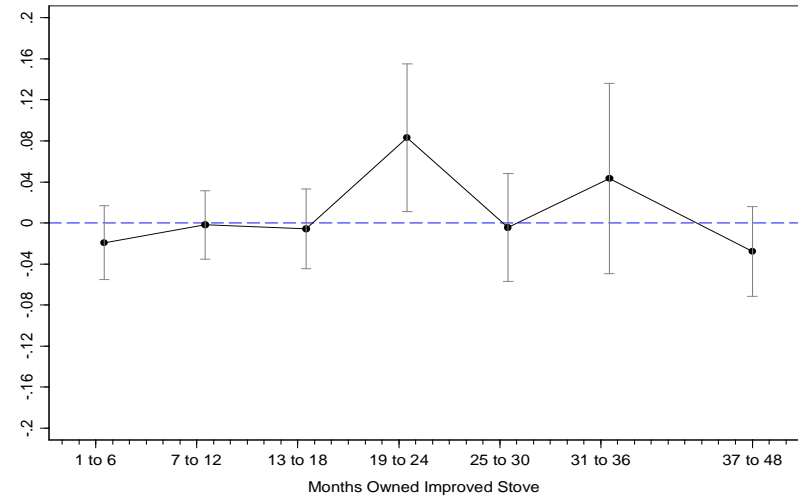
Panel B: Cough in the Last 30 Days
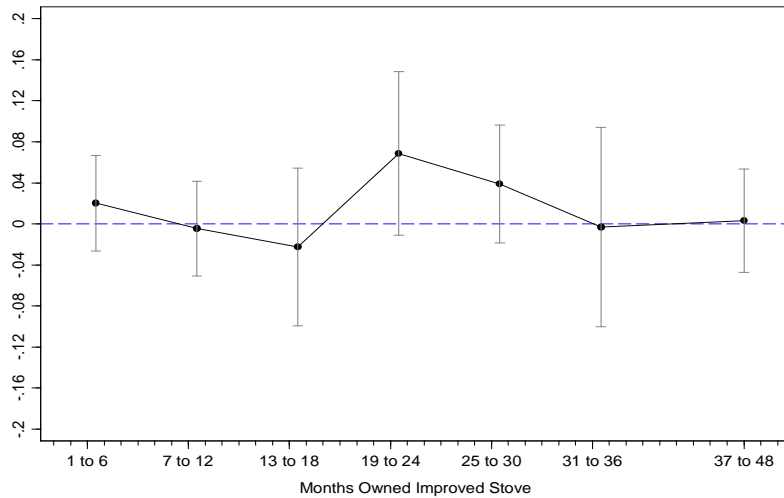
Panel D: Any Illness in Last 30 Days
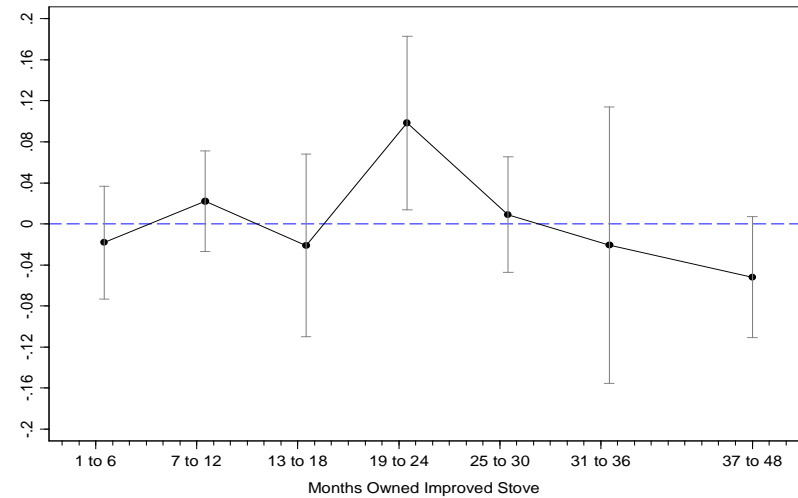
Notes: These figures show the difference in health outcomes between the treatment and control groups, by months since stove construction in the village, conditional on village x month of survey x year fixed effects. The bars represent the 95th percent confidence interval. Lottery 1 and Lottery 2 are grouped together as treatment. For continuous variables, the top 1 percent of values are dropped. BMI is standardized using values from the 2000 US CDC Population of Children.

Figure 5C: Reduced Form Effect of Stove Offer of the Health Outcomes for Children Aged 5 and Under in the Baseline, by Time
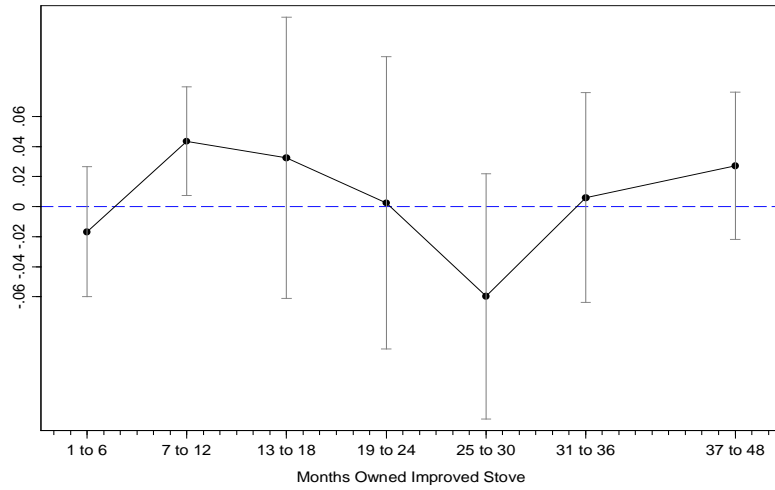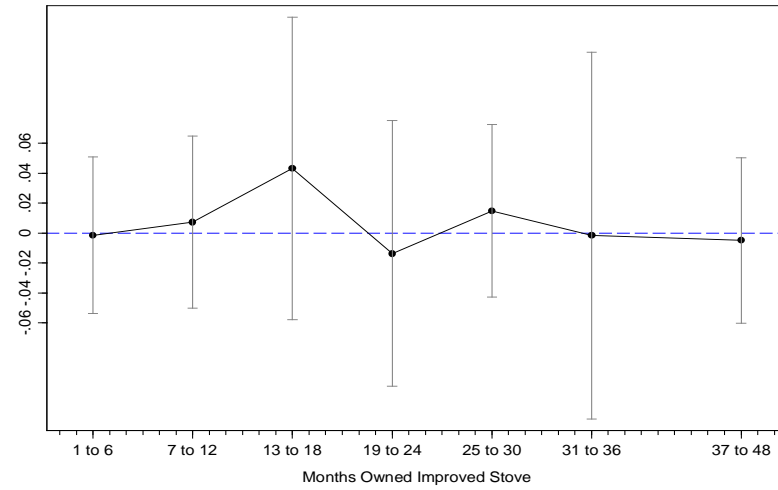

Notes: These figures show the difference in health outcomes between the treatment and control groups, by months since stove construction in the village, conditional on village x month of survey x year fixed effects. The bars represent the 95th percent confidence interval. Lottery 1 and Lottery 2 are grouped together as treatment. For continuous variables, the top 1 percent of values are dropped. BMI is standardized using values from the 2000 US CDC Population of Children.

A. Data Collection

In this section, we provide a comprehensive description of our data collection processes. As shown in Appendix Figure 3, we collected baseline data from January to July 2006, prior to the rollout of the intervention.[1] The baseline survey consisted of several modules. First, we conducted a comprehensive household module to collect data on household composition (size, as well as each member's age, sex, and relationship to the household head), demographics (education levels, caste, religion), economic indicators (asset ownership, indebtedness), and consumption patterns. As part of this module, we also collected information on the households' stove types, stove usage, housing construction, and fuel use. After the household module was administered, each member of the family was individually interviewed. For each child under the age of 14, his or her primary care-giver answered a child module on his or her behalf. In this module, we collected detailed data on stove use, such as how many meals an individual cooked that week, as well as a series of recall questions that were designed to gauge both respiratory and general health. For example, we collected data on whether one had a cough in the last 30 days, whether one had a fever in the last 30 days, and health expenditures. In addition, to understand the relationship between indoor air pollution and productivity, we collected detailed data on employment status and time-use patterns for adults over the last 24-hours, and school enrollment and attendance for children.

In the third, and last, component of the baseline survey, a specially trained enumeration team conducted a physical health check for each household member (both adults and children).[2] The examination included detailed biometric measurements, such as height, weight, and arm circumference. Most importantly, the examination included two tests designed to gauge exposure to smoke and respiratory functioning. First, to gauge smoke exposure, the team measured carbon monoxide (CO) in exhaled breath with a Micro Medical CO monitor.[3] CO is a

---

[1] From June to November 2007, we conducted the baseline for the five villages that we added to the study.

[2] The baseline survey was conducted by an outside survey company, while the baseline health checks were conducted by an internally hired and trained team. Therefore, the health checks were conducted on a different schedule as the main survey.

[3] Note that we did not measure ambient pollutants (either CO or PM). Ambient measures are less interesting to measure than exposure measures, as individuals may undertake fewer behaviors to protect themselves from smoke if ambient measures fall and could, in fact, end up with a higher level of exposure. If we conducted only ambient

biomarker of recent exposure to air pollution from biomass combustion, and therefore it can be used to proxy an individual's personal exposure to smoke from cooking stoves.

Second, we administered spirometry tests, which are designed to gauge respiratory health by measuring how much air the lungs can hold and how well the respiratory system can move air in and out of the lungs.[4] The tests were conducted using guidelines from both the equipment directions and American Association for Respiratory Care.[5]

After the initial lotteries were conducted and the stoves were built, we conducted the Continuous Health Survey (CHS) from January 2007 to June 2008. Each household was visited once. The survey consisted of two key modules. The first module was a household survey, in which we collected data on whether individuals used the new stoves, whether the stoves were used properly, whether individuals cleaned the stoves, and questions to gauge user satisfaction with the stoves. This module also included recall questions on health for each household member. Finally, we collected information on employment and school attendance in the last 30 days to gauge the impact of the respiratory health on productivity. The second module consisted of the same physical health exam used in the baseline survey. However, for this survey, we only conducted the health exam for women and children.[6]

From May 2008 to December 2009, we conducted the midline survey. This survey replicated the baseline survey, collecting data on stove use, recall health, and productivity. As part of this survey, we also conducted the health examination for all household members (men, women, and children). Finally, the endline survey was conducted from February to December

---

measures we would see a decline, even though their actual exposure may not have decreased due to behavioral changes. We focused on CO, which has been argued to be a good proxy for PM. Collecting data on PM exposure is difficult in this setting: tubes must be attached to the subjects for 24 hours and the equipment requires controlled temperature, careful transferring of samples, and proper laboratories for testing. Given the conditions of rural Orissa, controlling the samples would be near impossible at such a large scale. However, McCracken and Smith (1998) report a strong correlation between the average concentrations of CO and $PM_{2.5}$ in the kitchen during water boiling tests. They conclude that this implies "the usefulness of CO measurements as an inexpensive way of estimating $PM_{2.5}$ concentrations," even if it is not an exact proxy (see Ezzati [2002] for a discussion of this).

[4] In contrast to peak flow tests, which are easier to administer, spirometry readings can be used to diagnose obstructive lung disorders (such as chronic obstructive pulmonary disease [COPD] and asthma), and also restrictive lung disorders. Further, they are the only way to obtain measurements of lung function that are comparable across individuals (Beers et al., 1999).

[5] A manual spirometer was used in the baseline, continuous health survey (CHS), and a portion of the midline. The enumerators would take up to seven readings for each individual, until there were at least three satisfactory readings and at least two FEV1 readings within 100mL or 5 percent of each other. Electronic spirometers were adopted halfway through the midline. The new machines indicated when satisfactory readings had been completed and saved the best reading for each individual.

[6] Given that we were only able to visit each household once and that men are more difficult to find at home, including men would have been a challenge due to attrition.

2010. The endline survey was very similar to the baseline and midline, covering most of the questions that were in these surveys, with some modifications based on our experiences in the last two surveys. Specifically, we shortened the time use and household sections, and we added more comprehensive questions on infant mortality and pregnancy outcomes. Most importantly, we added a section on cooking and maintenance practices with the improved stoves (e.g., do households clean the stove, do households cover both pots when cooking, do households clean the chimney) and beliefs about the stove (e.g., whether they used more fuel, why they would recommend it).

In addition to the main surveys, we conducted village sweeps from May 2007 to January 2010 to collect data on recent births. Specifically, we tracked pregnancies that occurred during the scope of the study and then followed up on birth outcomes. However, we were often prevented from gathering data soon after each birth because many women relocate to their parents' home in the late stages of pregnancy and road conditions were poor during periods of monsoons. Therefore, we suspended the pregnancy sweeps before the endline survey (that started in February 2010) and instead collected the rest of these data during the endline survey.

Besides the large-scale surveys, we conducted a number of shorter surveys to assess stove ownership, repairs, and costs more frequently. In March and July of 2007, we conducted our first sweeps to assess which households built a stove after Lottery 1 and whether that stove was in good condition. We conducted these surveys and the two other similar stove surveys at the same time as the CHS. The first Chulha Monitoring Survey (CMS) was administered from August 2007 to January 2008 and the first Stove Survey (SS) was conducted in April 2008. Both surveys asked households if they had built an improved stove, why they did not build one if they had not, and the current condition of the stove if they had. The CMS additionally obtained information on whether stove owners used their stove properly and their perception of fuel and time efficiency of the stove compared to the traditional stove. The second round of the CMS overlapped with the midline survey administration and was conducted between March and December 2009. Finally, the second round of the SS was completed between January and May 2010, during the beginning of the endline survey administration.

Note that we also meticulously collected data on stove breakages, repairs and costs for both traditional and improved stoves in a stove cost survey that was administered concurrently with both the midline and endline surveys.[7]

Lastly, throughout the project we also collected administrative data on the functioning of this Gram Vikas program. Specifically, we collected data on lottery participation and the lottery outcomes in each village.

B. Experimental Validity: Details

In this section, we provide a more detailed description of experimental validity tests. Appendix Table 2A and 2B provide a test of the randomization for the baseline demographics, stove and fuel use, and health for those identified as primary cooks and children in the baseline. In Columns 1, 2, and 3 of both tables, we present the mean of each variable for the households that were assigned to the pure control group (those who never won a lottery), Lottery 1 winners, and Lottery 2 winners, respectively. Standard deviations are listed below the means in parentheses. The difference in means between the Lottery 1 winners and the control group are presented in Column 4, differences between Lottery 2 winners and the control group are listed in Column 5, and differences between Lottery 1 and Lottery 2 winners are presented in Column 6, conditional on village fixed effects. Standard errors that are clustered at the household level are shown in parentheses in Columns 4–6. In the final row of each panel we provide the p-value for a test of joint significance of the difference across each outcome variable.

The treatment groups appear to be generally well-balanced across the 59 baseline characteristics that we consider. Out of the 177 differences in Columns 4–6 of both tables, only 19 (or 10 percent) are significant at the 10 percent level or more, as would be predicted by chance. There are some notable differences in health characteristics: difference in any illness for primary cooks between Lottery 2 and the control (Column 5) and Lottery 1 and Lottery 2 (Column 6), as well as differences in whether a child had a cough in the last 30 days between Lottery 1 and both the control (Column 1) and Lottery 2 (Column 6). However, in testing the joint significance of health differences for primary cooks and children, only one of the set of six

---

[7] We additionally administered a cost survey to shopkeepers to collect information on the market price, product life, and repair costs for other clean stoves (i.e., LPG, mini-LPG, kerosene, and electric heaters). The addition of the Gram Vikas cost survey in the endline provides additional data on the price of different pipes used for constructing the improved stoves, and how much each household pays construction workers for their services.

differences is significant at the 10 percent level or more (the difference between Lottery 1 and Lottery 2 for primary cooks). Despite the fact that the data appear well-balanced, to increase precision in the ensuing analysis on the effect of the stoves on CO exposure and health, we will condition the regressions on the baseline values (although note that, in practice, controlling for the baseline values does not alter the magnitude nor significance of our estimates).[8]

In Appendix Table 3, we test for differential attrition in each of the four main surveys (baseline, CHS, midline, and endline). In each of the odd columns, we present the coefficient estimates when we regress a dummy variable for survey attrition on the treatment dummy. In the even columns, we present coefficient estimates from regressing an interaction of the treatment dummy with a set of indicator variables for survey round. In Columns 1 and 2, we study household attrition, while in the remaining columns we test for attrition within the sample of individuals. All regressions are clustered at the household level and include village by survey fixed effects. On average, we do not observe a significant difference in survey attrition for households (Column 1) and note that the magnitude of this coefficient estimate is also near zero. Further, we do not observe significant differences across any of the four main survey waves (Column 2). For individuals (females, primary cooks and children), we only observe a small, significant difference in attrition for primary cooks in the CHS and children in the baseline (significant at the 10 percent level). In total, out of the twenty differences that we explore in this table only two are significant at the 10 percent level, which is what one would expect by chance. These findings suggest that differential attrition is not a source of bias in the analysis.

---

[8] The results for women are similar to those of primary cooks. Out of the 39 differences (13 variables times three differences) we explore for women, only four are significant (or 10 percent, which is what would be predicted by chance). We omit them for brevity, as our subsequent health analysis will be primarily centered on primary cooks. However, they are available upon request.

Appendix Table 1: Correlations between the Number of Meals Cooked during Last Week with a Clean Stove and Baseline Smoke Exposure and Health (NO CONTROLS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)

Panel A: Females

	CO	FEV1	FEV1/FVC * 100	BMI	Cough or Cold	Sore Eyes	Headache	Phlegm	Wheeze	Tight Chest	Any Illness	Health Expenditures
Meals	-0.049	0.003*	0.058*	0.080***	-0.006***	-0.006***	0.002	-0.002	-0.001***	-0.001*	0.001	2.202**
	(0.030)	(0.002)	(0.031)	(0.013)	(0.002)	(0.002)	(0.002)	(0.001)	(0.000)	(0.001)	(0.001)	(0.891)
N	3,283	2,681	2,614	3,645	4,159	4,160	4,158	4,149	4,149	4,149	4,161	4,105

Panel B: Primary Cooks

	CO	FEV1	FEV1/FVC * 100	BMI	Cough or Cold	Sore Eyes	Headache	Phlegm	Wheeze	Tight Chest	Any Illness	Health Expenditures
Meals	-0.044	0.002	0.051	0.099***	-0.006**	-0.008***	0.005*	-0.001	-0.001***	-0.001*	0.000	2.650**
	(0.036)	(0.003)	(0.039)	(0.016)	(0.003)	(0.002)	(0.003)	(0.002)	(0.000)	(0.001)	(0.002)	(1.164)
N	1,967	1,654	1,617	2,092	2,421	2,421	2,421	2,420	2,420	2,420	2,421	2,393

Panel C: Children Aged 13 and Under in the Baseline

	CO	BMI	Cough	Consult for Fever	Earache	Skin Infection	Vomit	Weakness	Abdominal Pain	Hearing Problems	Vision Problems	Worms	Diarrhea	Any Illness	Health Expenditures
Meals	-0.116***	0.011	-0.002	-0.003	0.001	-0.000	-0.001	0.000	-0.001	-0.000	0.001	0.003	-0.001	-0.002	-0.053
	(0.043)	(0.009)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.000)	(0.001)	(0.002)	(0.001)	(0.003)	(0.529)
N	507	2,659	3,293	3,232	3,293	3,292	3,293	3,292	3,292	3,293	3,293	3,289	3,293	3,293	3,199

Panel D: Children Aged 5 and Under in the Baseline

		BMI	Cough	Consult for Fever	Earache	Skin Infection	Vomit	Weakness	Abdominal Pain	Hearing Problems	Vision Problems	Worms	Diarrhea	Any Illness	Health Expenditures
Meals		-0.001	-0.004	-0.005	0.001	-0.001	0.000	-0.002	-0.002	-0.001***	0.002	0.001	-0.001	-0.001	0.651
		(0.015)	(0.004)	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.003)	(0.000)	(0.002)	(0.003)	(0.002)	(0.004)	(0.927)
N		915	1,379	1,353	1,379	1,378	1,379	1,378	1,378	1,379	1,379	1,376	1,379	1,379	1,336

Notes: This table provides the correlation between the number of meals cooked with a clean stove at time of baseline and each variable listed in the table.   All regressions are estimated using OLS, include village x month of survey x year of survey fixed effects, and standard errors are clustered at the household level. For continuous variables, the top 1 percent of values are dropped. BMI for children is standardized using values from the 2000 US CDC Population of Children. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2A: Randomization Check for Baseline Demographic Characteristics and Stove Use

	Means			Differences, Conditional on Village FE		
	Control	Lottery1	Lottery2	Lottery1 - Control	Lottery2 - Control	Lottery1 - Lottery2
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Demographics						
Household Size	6.41	6.73	6.58	0.3548**	0.2979*	0.1802
	(3.18)	(3.78)	(3.52)	(0.1583)	(0.1629)	(0.1798)
Monthly Per Capita Household Expenditures	470.10	475.05	483.36	9.5308	16.3650	-10.3258
	(295.12)	(306.29)	(296.15)	(13.8931)	(14.5893)	(14.8053)
Minority Household (Scheduled Caste or Tribe)	0.40	0.42	0.48	-0.0225**	0.0006	-0.0241**
	(0.49)	(0.49)	(0.50)	(0.0113)	(0.0120)	(0.0122)
Has Electricity in Household	0.47	0.48	0.46	0.0262	0.0180	0.0102
	(0.50)	(0.50)	(0.50)	(0.0204)	(0.0215)	(0.0213)
Male Head Ever Attended School	0.71	0.71	0.66	0.0097	-0.0244	0.0272
	(0.45)	(0.45)	(0.47)	(0.0218)	(0.0243)	(0.0241)
Male Head Literate	0.62	0.59	0.54	-0.0171	-0.0433	0.0190
	(0.49)	(0.49)	(0.50)	(0.0250)	(0.0267)	(0.0269)
Female Head Ever Attended School	0.31	0.34	0.30	0.0406*	-0.0022	0.0311
	(0.46)	(0.47)	(0.46)	(0.0225)	(0.0234)	(0.0234)
Female Head Literate	0.21	0.21	0.18	0.0079	-0.0201	0.0157
	(0.41)	(0.41)	(0.39)	(0.0201)	(0.0207)	(0.0206)
Female Has a Savings Account	0.64	0.69	0.69	0.0356	0.0304	0.0033
	(0.48)	(0.46)	(0.46)	(0.0224)	(0.0233)	(0.0230)
P-value from Joint Test				0.09	0.80	0.30
Panel B: Baseline Stove Characteristics and Fuel Use						
Traditional Stove	0.99	1.00	0.99	0.0047	-0.0014	0.0051
	0.10	0.07	0.10	(0.0042)	(0.0052)	(0.0044)
Any Type of "Clean Stove"	0.23	0.24	0.23	0.0176	0.0041	0.0028
	0.42	0.43	0.42	(0.0192)	(0.0199)	(0.0199)
Improved Stove	0.01	0.01	0.01	-0.0012	-0.0026	0.0038
	(0.11)	(0.10)	(0.07)	(0.0051)	(0.0049)	(0.0044)
Kerosene	0.11	0.10	0.10	-0.0034	-0.0156	0.0080
	(0.31)	(0.31)	(0.29)	(0.0148)	(0.0153)	(0.0150)
Biogas	0.03	0.03	0.03	-0.0045	-0.0037	-0.0049
	(0.17)	(0.16)	(0.17)	(0.0070)	(0.0074)	(0.0072)
LPG	0.03	0.05	0.05	0.0124	0.0116	0.0030
	(0.18)	(0.21)	(0.22)	(0.0094)	(0.0101)	(0.0109)
Electric	0.10	0.12	0.11	0.0270**	0.0144	0.0079
	(0.29)	(0.32)	(0.31)	(0.0137)	(0.0141)	(0.0147)
Coal	0.00	0.00	0.01	-0.0022	0.0002	-0.0025
	(0.07)	(0.05)	(0.07)	(0.0029)	(0.0036)	(0.0032)
Cooked Most Meals with Traditional Stove in Last Week	0.94	0.93	0.92	-0.0160	-0.0239*	0.0085
	(0.23)	(0.26)	(0.28)	(0.0115)	(0.0125)	(0.0133)
Meals Cooked Last Week	13.65	13.82	13.80	0.1659	0.2457	-0.0444
	(4.34)	(4.02)	(3.77)	(0.2018)	(0.2069)	(0.1977)
Meals Cooked Last Week with Traditional Stove	12.63	12.59	12.59	-0.0142	0.0955	-0.1087
	(4.83)	(4.54)	(4.57)	(0.2217)	(0.2336)	(0.2265)
% Primary Cook Female	0.32	0.31	0.32	-0.0115	-0.0062	-0.0065
	(0.47)	(0.46)	(0.47)	(0.0133)	(0.0142)	(0.0139)
Meals Cooked in Open Area Last Week	7.40	7.28	7.72	0.0744	-0.0465	0.1195
	(7.06)	(7.17)	(7.31)	(0.3140)	(0.3254)	(0.3255)
Meals Cooked in Semi-open Area Last Week	5.06	5.18	5.03	0.0956	-0.0781	0.1429
	(7.03)	(6.84)	(6.68)	(0.1551)	(0.1578)	(0.1605)
Meals Cooked in Enclosed Area Last Week	0.83	0.94	0.79	0.0091	0.0023	0.0090*
	(3.19)	(3.27)	(3.05)	(0.0057)	(0.0066)	(0.0053)
Ever Use Wood	0.98	0.99	0.99	0.4617	0.3894	-0.3731
	(0.14)	(0.09)	(0.12)	(0.4799)	(0.4716)	(0.5133)
Minutes Spent Gathering Wood Yesterday (if gathered wood)	4.96	5.10	4.80	-0.0204	-0.0561	-0.0551
	(6.13)	(7.04)	(5.57)	(0.3554)	(0.4300)	(0.4002)
Wood Used for Last Meal (in kg)	5.19	5.04	4.88	0.0369**	-0.0170	0.0541***
	(7.23)	(6.06)	(8.38)	(0.0162)	(0.0178)	(0.0170)
Meals Per Bundle of Wood	0.82	0.87	0.81	-0.0195	0.0117	-0.0199
	(0.39)	(0.34)	(0.39)	(0.0195)	(0.0202)	(0.0203)
Household Gathers Wood	0.37	0.33	0.36	-0.0117	-0.0100	-0.0015
	(0.48)	(0.47)	(0.48)	(0.0139)	(0.0146)	(0.0142)
Ever Bought Wood	0.20	0.19	0.21	-3.9369	-6.7556	2.6789
	(0.40)	(0.39)	(0.41)	(5.8293)	(6.0798)	(5.0492)
Ever Sold Wood	37.24	34.32	30.63	-0.0115	-0.0062	-0.0065
	(145.87)	(122.90)	(123.91)	(0.0133)	(0.0142)	(0.0139)
P-values from Joint Test				0.21	0.96	0.28

Appendix Table 2B: Randomization Check for Baseline CO and Health

	Means			Differences, Conditional on Village FE		
	Control	Lottery1	Lottery2	Lottery1 - Control	Lottery2 - Control	Lottery1 - Lottery2
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Primary Cooks			
CO	7.93	7.52	7.82	-0.4378	0.0508	-0.5374
	(6.62)	(5.60)	(6.36)	(0.3296)	(0.3742)	(0.3305)
FEV1	1.97	1.97	1.96	0.0093	0.0022	0.0013
	(0.37)	(0.37)	(0.37)	(0.0220)	(0.0232)	(0.0228)
FVC	2.30	2.32	2.29	0.0218	0.0079	0.0090
	(0.45)	(0.43)	(0.42)	(0.0262)	(0.0272)	(0.0263)
FEV1/FVC	89.54	89.63	89.72	-0.0514	-0.2857	0.1048
	(5.93)	(6.49)	(5.88)	(0.3587)	(0.3650)	(0.3832)
BMI	19.06	18.76	18.89	-0.2063*	0.0219	-0.2658**
	(2.49)	(2.54)	(2.51)	(0.1250)	(0.1343)	(0.1308)
Cold or Cough	0.53	0.48	0.54	-0.0420*	0.0197	-0.0546**
	(0.50)	(0.50)	(0.50)	(0.0242)	(0.0253)	(0.0250)
Any Illness	0.85	0.85	0.91	-0.0021	0.0549***	-0.0574***
	(0.36)	(0.36)	(0.29)	(0.0176)	(0.0168)	(0.0166)
Phlegm	0.14	0.12	0.14	-0.0200	-0.0006	-0.0116
	(0.35)	(0.33)	(0.34)	(0.0163)	(0.0177)	(0.0170)
Headache	0.50	0.48	0.49	-0.0286	-0.0198	-0.0048
	(0.50)	(0.50)	(0.50)	(0.0238)	(0.0251)	(0.0252)
Sore eyes	0.29	0.28	0.28	-0.0058	0.0124	-0.0098
	(0.45)	(0.45)	(0.45)	(0.0216)	(0.0227)	(0.0227)
Wheezing	0.01	0.01	0.01	-0.0029	-0.0013	-0.0029
	(0.12)	(0.10)	(0.10)	(0.0052)	(0.0054)	(0.0049)
Tightness in Chest	0.04	0.04	0.05	0.0029	0.0032	-0.0056
	(0.20)	(0.21)	(0.21)	(0.0100)	(0.0106)	(0.0103)
Total Health Expenditures	56.62	64.69	64.00	5.9421	2.0802	5.1422
	(164.67)	(166.53)	(179.46)	(8.0816)	(9.2103)	(9.1846)
P-value from Joint Test				0.221	0.390	0.037
			Panel B: Children Under Aged 13			
CO	6.63	5.90	6.89	-0.4398	0.2126	-0.7463
	(5.48)	(4.74)	(5.55)	(0.5464)	(0.7118)	(0.6406)
BMI	-1.82	-1.89	-1.85	-0.0479	0.0342	-0.0772
	(1.30)	(1.30)	(1.27)	(0.0704)	(0.0732)	(0.0707)
Cough	0.39	0.42	0.38	0.0517**	-0.0026	0.0590**
	(0.49)	(0.49)	(0.49)	(0.0239)	(0.0255)	(0.0243)
Consulted for Fever	0.28	0.25	0.28	-0.0256	-0.0204	-0.0137
	(0.45)	(0.43)	(0.45)	(0.0202)	(0.0221)	(0.0212)
Earache	0.09	0.09	0.09	0.0102	0.0024	0.0020
	(0.28)	(0.29)	(0.28)	(0.0130)	(0.0139)	(0.0140)
Skin	0.13	0.14	0.13	0.0120	0.0009	0.0072
	(0.34)	(0.34)	(0.34)	(0.0164)	(0.0168)	(0.0164)
Any Illness	0.73	0.73	0.74	-0.0007	-0.0102	-0.0032
	(0.44)	(0.44)	(0.44)	(0.0222)	(0.0227)	(0.0222)
Vision Problems	0.01	0.01	0.01	-0.0038	-0.0004	-0.0011
	(0.12)	(0.10)	(0.11)	(0.0046)	(0.0046)	(0.0046)
Hearing Problems	0.01	0.02	0.01	0.0031	-0.0034	0.0094*
	(0.11)	(0.12)	(0.10)	(0.0049)	(0.0047)	(0.0051)
Vomiting	0.07	0.08	0.09	0.0094	0.0116	-0.0026
	(0.25)	(0.27)	(0.28)	(0.0115)	(0.0124)	(0.0129)
Diarrhea	0.07	0.08	0.08	-0.0022	-0.0082	0.0091
	(0.26)	(0.27)	(0.27)	(0.0118)	(0.0129)	(0.0124)
Abdominal Pain	0.14	0.15	0.14	0.0039	0.0008	0.0078
	(0.35)	(0.35)	(0.35)	(0.0162)	(0.0172)	(0.0166)
Worms	0.09	0.11	0.09	0.0165	-0.0108	0.0209
	(0.28)	(0.31)	(0.28)	(0.0144)	(0.0158)	(0.0157)
Weakness	0.21	0.21	0.22	-0.0100	-0.0062	-0.0066
	(0.41)	(0.41)	(0.42)	(0.0192)	(0.0215)	(0.0203)
Total Health Expenditures	44.61	48.89	45.26	4.9607	0.7712	3.7218
	(101.22)	(99.72)	(106.05)	(4.6885)	(5.2102)	(5.0337)
				0.777	0.834	0.775

Appendix Table 3:   Testing for Survey Attrition

	Household		Females		Primary Cooks		Children	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat	0.000		-0.006		0.001		0.013	
	(0.005)		(0.008)		(0.008)		(0.010)	
Treat * Baseline		0.003		-0.004		-0.001		0.031*
		(0.006)		(0.014)		(0.003)		(0.017)
Treat * CHS		0.008		0.005		0.026*		0.004
		(0.008)		(0.013)		(0.014)		(0.013)
Treat * Midline		-0.007		-0.014		-0.011		0.007
		(0.009)		(0.012)		(0.016)		(0.017)
Treat * Endline		-0.003		-0.010		-0.011		0.010
		(0.010)		(0.009)		(0.017)		(0.016)
N	10,300	10,300	27,776	27,776	10,040	10,040	25,800	25,800

Notes: This table provides results on whether there exists differential survey attrition by treatment status.   The dependent variable is a dummy variable that indicates whether the household (or individual) was not included in the survey. CHS is the continuous health survey conducted between the baseline and midline survey. All regressions include village x survey fixed effects and standard errors are clustered at the household level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix Table 4: Reasons for Not Having a Stove, by Year of Stove Being Offered in Your Village

	Year 1 (1)	Year 2 (2)	Year 3 (3)	Year 4 (4)
Insufficient Kitchen Space/Family Size and Stove Do Not Match	0.28	0.34	0.22	0.19
Does Not Want A Double Pot	0.01	0.02	0.02	0.01
Already Owns a Better Stove	0.06	0.08	0.03	0.04
Will Build Soon	0.27	0.21	0.38	0.15
Not Interested in Building	0.07	0.15	0.06	0.26
Destroyed by User	0.02	0.06	0.23	0.32
Other	0.24	0.11	0.06	0.04

Notes: This table provides information on why households did not have a stove, by years since stove was offered in their village.

Appendix Table 5: Reduced Form Effect of Stoves on Male Smoking

	Male Smokes
	(1)
Panel A: Overall Treatment Effect	
Treat	0.009
	(0.009)
Panel B: By Months Since Stove Construction	
Treat x I(0 to 12 mo)	0.008
	(0.014)
Treat x I(13 to 24 mo)	-0.013
	(0.021)
Treat x I(25 to 36 mo)	0.015
	(0.012)
Treat x I(37 to 48 mo)	0.014
	(0.015)
N	5,874
Control Group Mean	0.121

Appendix Figure 1: Traditional and Gram Vikas Improved Stoves
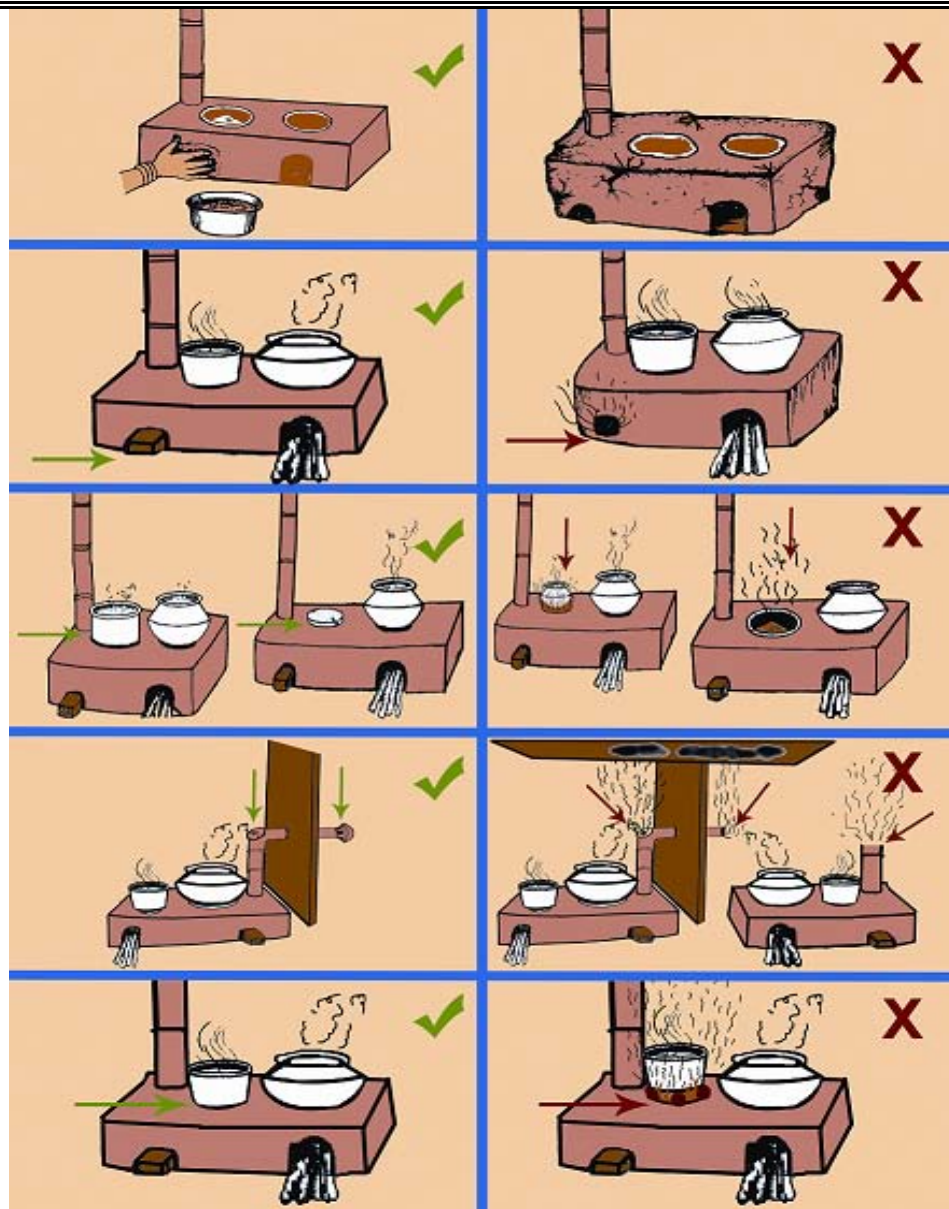
Panel A: Traditional Stove

Panel B: Gram Vikas Improved Stove

Appendix Figure 2: Example of Gram Vikas Training Material

Appendix Figure 3: Timeline of Major Surveys and Construction

2006	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Baseline	X	X	X	X	X	X	X					
Lottery 1 Stove Construction									X	X	X	X
CHS												
Pregnancy Survey					X	X	X	X	X	X	X	X
Midline												
Lottery 2 Stove Construction												
Endline												

2007	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Baseline						X	X	X	X	X	X	
Lottery 1 Stove Construction	X	X	X							X	X	X
CHS								X	X	X	X	X
Pregnancy Survey	X	X	X	X	X	X	X	X	X	X	X	X
Midline												
Lottery 2 Stove Construction												
Endline												

2008	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Baseline												
Lottery 1 Stove Construction												
CHS	X	X	X	X	X	X						
Pregnancy Survey	X	X	X	X	X	X	X	X	X	X	X	X
Midline					X	X	X	X	X	X	X	X
Lottery 2 Stove Construction												
Endline												

2009	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Baseline												
Lottery 1 Stove Construction												
CHS												
Pregnancy Survey	X	X	X	X	X	X	X	X	X	X	X	X
Midline	X	X	X	X	X	X	X	X	X	X	X	X
Lottery 2 Stove Construction					X	X	X	X	X	X	X	X
Endline												

2010	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Baseline												
Lottery 1 Stove Construction												
CHS												
Pregnancy Survey	X											
Midline												
Lottery 2 Stove Construction	X	X	X	X								
Endline		X	X	X	X	X	X	X	X			

Notes: This table provides a timeline of each major survey we conducted during the course of the project, as well as a timeline of the lotteries and intervention. CHS is the continuous health survey that we conducted between the baseline and midline surveys.

Appendix Figure 4: Household Sample Sizes, By Survey

Panel A: Primary Household and Health Surveys

Baseline Survey 2575 HH	➡	CHS 2427 HH	➡	ML: 2400 HH [ML Stove Cost: 2212]	➡	EL: 2048 HH [EL Stove Cost: 2030]
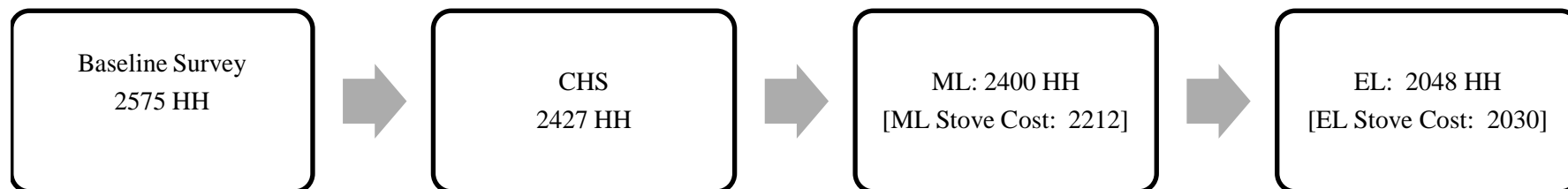
Panel B: Primary Surveys on the Condition of the Stove

| Stove Building Survey 2575 HH	➡	Stove Monitoring Survey I 2281 HH	➡	Stove Status I 2474 HH	➡	Stove Monitoring Survey II 2377 HH	➡	Stove Status II 2423 HH

Notes: This table documents the majority of surveys that were conducted during the course of the study. HH is the number of participating households. CHS is the continuous health survey that we conducted between the baseline and midline surveys. ML is the midline survey, while EL is the endline survey. We conducted the stove costs survey in conjunction with the midline and endline. Stove status surveys provide basic information on the stoves, while stove monitoring is a more in-depth survey of use.

Appendix Figure 5: Ever Had a Stove Breakage
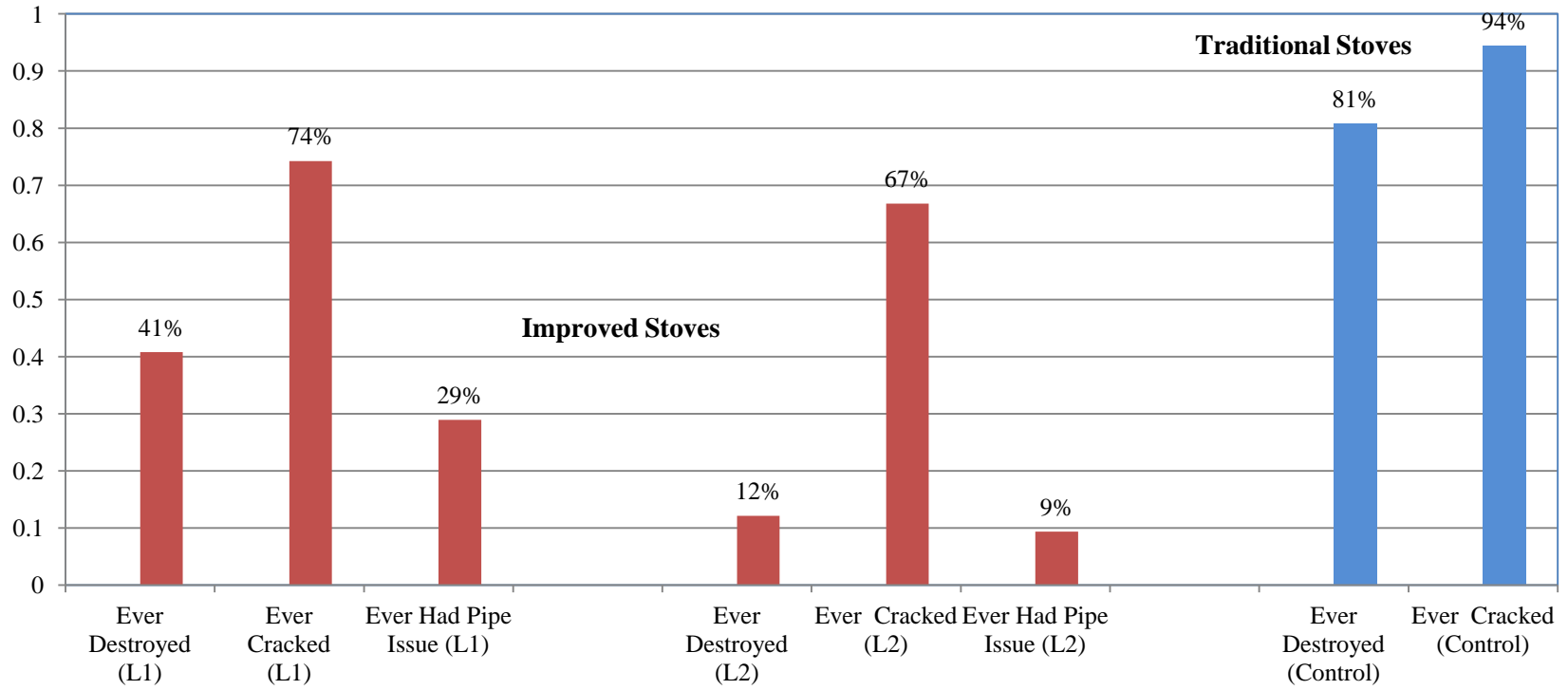
Appendix Figure 6: Beliefs on Stove Quality
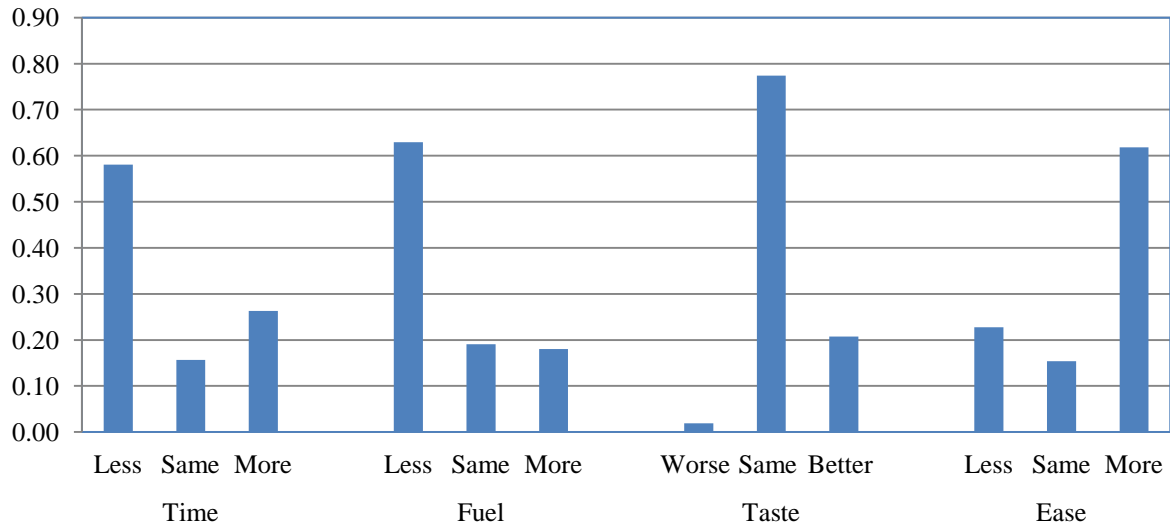
Notes: This table provide sample statistics on self-reported beliefs with the GV improved cooking stoves for those who own one.