The Effects of Fuel-Efficient Cookstoves on Fuel Use, Particulate Matter, and Cooking Practices: Results from a Randomized Trial in Rural Uganda

By THERESA BELTRAMO, GARRICK BLALOCK, STEPHEN HARRELL, DAVID I. LEVINE, AND ANDREW M. SIMONS*

Smoky cookfires contribute to global climate change and kill approximately four million people annually. While many studies have examined the effects of fuel-efficient cookstoves, this study was among the first to do so while selling stoves at market prices. After introducing a fuel-efficient cookstove, fuelwood use and household air particulates declined by 12% and by smaller percentages after adjusting for observer-induced bias, or the Hawthorne effect. These reductions were less than laboratory predictions and fell well short of World Health Organization pollution targets. Even when introducing a second stove, most households continued to use their traditional stoves for most cooking.

Keywords: technology adoption; household air pollution; biomass fuel; climate change; Hawthorne effect; sensors

JEL codes: O10; O12; O13; Q53; Q56; I15

*Beltramo: United Nations High Commissioner for Refugees, Rue de Montbrilliant 92, Geneva 1201, Switzerland (e-mail: beltramo@unhcr.org); Blalock: Charles H. Dyson School of Applied Economics and Management, Cornell University, Ithaca, NY 14853 (e-mail: garrick.blalock@cornell.edu); Harrell: Department of Agriculture and Resource Economics, University of California, Berkeley, CA 94720 (email: stephenharrell@gmail.com); Levine: Walter A. Haas School of Business, University of California, Berkeley, 2220 Piedmont Ave., Berkeley, CA 94720 (e-mail:

levine@haas.berkeley.edu); Simons (corresponding author): Department of Economics, Fordham University, 441 E. Fordham Rd., Bronx, NY 10458 (e-mail: asimons5@fordham.edu).

I. Introduction

Almost 3 billion people cook with wood, charcoal, and dung using traditional cookstoves (Bonjour et al. 2013). These stoves cause environmental degradation (Bailis et al. 2015), global climate change (Ramanathan and Carmichael 2008), and an estimated four million deaths per year (Lim et al. 2012). Truly safe cooking likely requires clean fuels such as gas or electricity. Unfortunately, most people who cook with solid fuel lack an affordable and consistent supply of gas or electricity (Lewis and Pattanayak 2012; Rehfuess et al. 2010). In the short to medium term, fuel-efficient cookstoves that use less solid fuel than traditional stoves may reduce these environmental and health problems.

We experimentally examined the effects of a fuel-efficient cookstove, the Envirofit G3300 wood-burning stove, on wood use, household air pollution, and cooking behaviors in rural Uganda. Our work builds on important antecedents and extends previous literature in three key ways: (1) households purchased the new stove at the market price; (2) we provided households with a second fuel-efficient stove to see if a second cooking surface would limit stove-stacking; and (3) we adjusted for observer-induced bias, or the Hawthorne effect.

The first studies to document the relationship of stove usage, household air pollution, and human health were conducted in Kenya (Ezzati and Kammen 2001, 2002; Ezzati, Saleh, and Kammen 2000) and Guatemala (Smith et al. 2006; Smith et al. 2011; Smith-Sivertsen et al. 2009). More recently, Hanna, Duflo, and Greenstone (2016) examined the link between stove usage and household air pollution in India and found reductions in smoke inhalation in the first year, but no changes over longer periods. They suggested that the fade-out was due to a lack of stove maintenance by users. Bensch and Peters (2015) examined a stove designed to reduce fuelwood consumption in rural Senegal and found reductions in fuelwood use, smoke emissions, and smoke-related disease symptoms. Pillarisetti et al. (2014) examined stove usage in a sample of pregnant women in India and found that users experimented with the fuel-efficient stove at first, but that the use of the new stove declined over time. Moreover, by one year after introduction, the sampled households used traditional stoves for 75% of their cooking.

Similar to the studies of Hanna, Duflo, and Greenstone (2016) and Bensch and Peters (2015), we measured stove use in the short term (a year or less) and over the long term (a 3.5 year follow-up). These two previous studies measured health outcomes (documented by medical personnel or self-reported). In contrast, we measured household level particulate matter (PM2.5) concentrations. Particulate matter concentrations have been directly linked to health problems in numerous studies (Chay and Greenstone 2003; Currie and Walker 2009; Smith-Sivertsen et al. 2009). Due to their small size (2.5 μg or less), these particles can reach deep into the lungs and are the best single indicator of risk for many respiratory-related diseases (Chowdhury et al. 2007).[1] Similar to Pillarisetti et al. (2014), we used unobtrusive temperature sensors to measure detailed household stove use over time.[2] However, unlike Pillarisetti et al. (2014), we introduced random variation in the assignment of when the stoves were delivered to causally examine the effects of the introduction of a fuel-efficient stove.

Our study extends previous literature in three important ways. First, we examine cooking behaviors among households that were willing to purchase the

[1]According to Pope III et al. (2002), each 10 $\mu g/m^3$ increase in long-term exposure to fine particulate matter is associated with approximately a 4%, 6%, and 8% increase in the risk of all-cause cardiopulmonary and lung cancer mortality, respectively.

[2]These stove usage monitors were pioneered by Ruiz-Mercado, Canuz, and Smith (2012).

new stove at market prices (and perhaps, therefore, value the stove more highly).[3] Because our results come from users who paid the market price for the fuel-efficient stove, our sample mimics those that would be most likely to purchase such a stove. There is a long-standing debate whether developing countries should charge for health improving products (latrines, mosquito bed nets, deworming medications, chlorine tablets, etc.) or if they should be distributed for free (Ashraf, Berry, and Shapiro 2010; Cohen and Dupas 2010; Dupas 2014; Fischer et al. 2019). A key part of this debate is the question of how usage of the product varies depending on the price paid. Generally, cookstoves have been given for free or highly subsidized in previous cookstove usage studies. Our study adds a new data point to quantify usage for users who paid market price for their cookstoves.

A second innovation in our study was that, after measuring stove usage when households had one fuel-efficient stove, we provided all households with a second fuel-efficient stove. Common cooking practice in the study area involved cooking with two pots simultaneously (e.g., rice and beans, or steaming bananas and cooking gravy). Stove stacking (the simultaneous use of the fuel-efficient stove and the traditional cooking technology) has been mentioned as a challenge to completely switching to fuel-efficient stoves (Masera, Saatkamp, and Kammen 2000; Pillarisetti et al. 2014; Ruiz-Mercado et al. 2011). This non-experimental

[3] Among these similar studies, Hanna, Duflo, and Greenstone (2016) distributed highly subsidized stoves (users paid US$0.75 for a US$12.50 stove), while Bensch and Peters (2015) and Pillarisetti et al. (2014) distributed stoves for free. Studies primarily focusing on the public health benefits of cookstoves typically distribute the cookstoves for free. For example, the randomized exposure study of pollution indoors and respiratory effects (RESPIRE) in Guatemala (Smith et al. 2006; Smith-Sivertsen et al. 2009), the Cooking and Pneumonia Study in Malawi (Mortimer et al. 2017), and the research on emissions, air quality, climate, and cooking technologies in Northern Ghana (REACCTING) study (Dickinson et al. 2015).

intervention allowed us to examine how important the lack of a second cooking surface was for continued use of the traditional stove.

A third innovation of our study was that we adjusted for observer-induced bias, or the Hawthorne effect. The Hawthorne effect has been mentioned as a potential source of bias in numerous cookstove studies (Bensch and Peters 2015; Ezzati, Saleh, and Kammen 2000; Pillarisetti et al. 2014; Smith-Sivertsen et al. 2009). By collecting sensor data both when observers were and were not present, we were able to measure and remove the source of this observer-induced bias.[4]

During the weeks when wood use and particulate matter were measured, we found that the randomized early introduction of the first fuel-efficient stove reduced wood use by 11.6% and particulate matter by 12.0%. Once both fuel-efficient stoves were introduced, wood use declined by 26.7% and particulate matter by 10.0%. However, we also found that participants cooked more on the fuel-efficient stoves and less on three-stone fires when observers were present, and that participants reversed these changes once observers left (Simons et al. 2017). When adjusting for this observer-induced bias, we found that the randomized early introduction of the first fuel-efficient stove may have only reduced wood use by 1.7% and particulate matter by 0.3%. Once both fuel-efficient stoves were introduced, after adjusting for the Hawthorne effect, we found wood usage may have declined by 2.5% compared to the baseline;

[4]This adjustment removes the bias from when observers were present compared to when no observers were present. We acknowledge that it is possible that the sensors themselves could have induced different behavior, however we feel that given the small size of the sensors (about the size of a coin) and the length of tracking (about six months) that the sensor was not salient enough to make a big difference in sustaining atypical cooking behaviors.

however, particulate matter may have increased[5] (an increase of 18.3% compared to the baseline).

Households used the new stoves more hours per day than the usage of the three-stone fires declined. The increase in total hours of stove usage blunted reductions in fuel use and household air pollution. At the same time, cooking on multiple surfaces most likely increased the utility of the cooks. It appears that cooks used each stove for the foods that fit it best. For example, low-heat simmering of rice, beans and unripe bananas was done on three-stone fires, and making sauces and boiling water for tea was done on the fuel-efficient stove. In the longer term (3.5 years), we found lower rates of disrepair than Hanna, Duflo, and Greenstone (2016).[6] Nevertheless, as in their study, we found low longer-term usage of the fuel-efficient stove.

Concerning related environmental problems, our findings suggest fuel-efficient cookstoves similar to the one used in our study and setting have, at best, marginal effects. The 12% reduction in fuel use (upon introduction of the first fuel-efficient stove) may generate small reductions in deforestation and carbon dioxide emissions, at least in the short term (though these reductions dissipated over the length of our study).

Concerning related health problems, the 12% reduction in particulate matter left the air 14 times more polluted than the World Health Organization (WHO) standard of 25 $\mu g/m^3$ (World Health Organization 2006). Thus, if clean air is a high priority, our findings suggest it is important to help consumers shift to safe

[5]Note that the introduction of the second fuel-efficient stove was not experimentally identified, and the difference in changes in particulate matter and wood use could have been due to a variety of factors, such as weather changes (i.e., wet wood burns less efficiently).

[6]This pattern makes sense as Hanna, Duflo, and Greenstone (2016) examined local artisan-built mud stoves, while the stoves used in our study were commercially manufactured from metal. The manufacturer (Envirofit Inc.) stated its stoves would last up to ten years. See https://envirofit.org/.

fuels such as gas or electricity and to find ways to encourage them to disable or move their smoky stoves outdoors.

II. Experimental Setting and Data

A. Background and Site and Stove Selection

We selected the Mbarara region of Uganda because it is rural, almost all families cooked on a traditional three-stone fire, households spent significant time gathering firewood or purchased firewood, and the local government was supportive of our work. In pre-experimental discussion groups, we confirmed that there was no active fuel-efficient cookstove intervention in the region, and that families spent significant time gathering wood (approximately 10–20 hours per week).

Most participants farm *matooke* (starchy cooking banana), potatoes, and millet and raise livestock. Prior to our experiment, almost all families cooked on a traditional three-stone fire (97%), usually located within a separate cooking hut. Most (62%) households had totally enclosed kitchens with no windows, while 38% had semi-enclosed kitchens with at least one window. Almost all cooking occurred in the detached cooking hut.

We implemented a series of companion studies in rural areas of the Mbarara District in southwestern Uganda from February to September 2012, focusing on the adoption of fuel-efficient stoves. These studies analyzed the household purchase decision, and they found that relieving liquidity constraints by allowing additional time for payments (Beltramo et al. 2015b) and providing a free trial with time payments allowed users to learn about the stoves' fuel savings properties (Levine et al. 2018) and greatly increased purchase rates (for example, from 5% to 57% in our setting in rural Uganda). We also examined how social networks affected purchasing (Beltramo et al. 2015a).

We marketed the Envirofit G3300 wood-burning stove, made by Envirofit International Inc. (Ft. Collins, CO, USA) (see Figure 1 for images of a traditional three-stone fire and the Envirofit G3300). This stove achieves relatively efficient fuel combustion by channeling airflow into the fire and directing heat upward through an insulated cylinder to the cooking surface. These design innovations allow fuel to burn at a controlled rate and enable more complete combustion than a three-stone fire. Emissions testing of the Envirofit G3300 in a controlled laboratory setting found average reductions in carbon monoxide (CO) of 65%, particulate matter reductions of 51%, and a reduction in fuel wood use of 50% compared with a three-stone fire (see Figure 2 for a copy of the emissions and performance report).

Before selecting the Envirofit G3300, we conducted a feasibility study that tested four different models of fuel-efficient stoves among households within the study zone.[7] The feasibility study included three focus groups and one town hall style meeting, which included a total of 85 participants. This study found that the participants preferred the Envirofit G3300. Additionally, during the feasibility study it was apparent that most households used two cooking points on most days. This finding informed our experimental design to distribute a second Envirofit to each household to give cooks the ability to completely substitute away from the use of traditional three stone fires.

B. Selection of Study Participants

In the first stage of the experiment, we randomly selected 12 parishes (units of government administration covering about 4,000–6,000 people), to receive a traditional full upfront payment sales offer and 14 parishes to receive a sales offer

[7]The full feasibility study report can be found here: https://www.cleancookingalliance.org/binary-data/CMP_CATALOG/file/000/000/153-1.pdf.

of a one-week free trial followed by four equal weekly time payments (see Levine et al. 2018). Within each parish, we recruited a local point person with the help of local government officials. We asked each focal point person to gather roughly 60 people together for a public sales meeting on a specified day. We did not tell the point person which sales offer his or her parish would receive.

At the sales meeting, participants completed a questionnaire that focused on household cooking and basic socioeconomic indicators. After this, the study team presented the Envirofit G3300, discussed the stove's features such as fuel savings and reduced pollution relative to traditional three-stone fires, gave a cooking demonstration, and presented the terms of the randomly selected sales offer. While the Envirofit was not commercially available in this region prior to our experiment, we sold it for the same retail price (40,000 Ugandan shillings [~US$16]) that it was selling for in parts of the country where it was available. We used the randomized assignment of the sales offer by parish as the identifying assumption, as used by Levine et al. (2018), to examine the barriers to purchase. In the current paper, to examine how often people used their stoves, our identification strategy was based on randomly assigning the timing of when purchasers received their Envirofit (we call them early buyers and late buyers). In each of the 14 parishes with the sales offer of a free trial plus time payment, we randomly selected 12 of the purchasing households for stove usage tracking. Therefore, all participants who had their stove usage tracked received the same sales offer at the extensive margin, and all participants fully paid for the stove according to the terms of the sales offer (one-week free trial, followed by four equal payments totaling 40,000 shillings).

Households were eligible to have their stove usage tracked if they mainly used wood as a fuel source, regularly cooked for eight or fewer persons (so that their cooking pots could fit on the Envirofit), someone was generally home every day, and cooking was largely done in an enclosed kitchen. In each parish, more than 12 households met these criteria and agreed to join the study; therefore, among those that agreed, we randomly selected 12 households per parish to track with the stove use monitors (SUMs). We then randomly assigned each of these 12 households to be an early buyer or late buyer. We asked both early and late buyers if they would agree to have SUMs immediately placed on their traditional three-stone fires (all agreed). We used the randomly assigned time of Envirofit delivery (early buyers vs. late buyers) as the identifying assumption for the causal claims made in this paper.

After participants consented to participate in the usage study, all existing three-stone fires were affixed with SUMs. Then, approximately four weeks after the SUM data collection began, the early buyers' group received their first Envirofit stove. Approximately four weeks after that, the late buyers received their first Envirofit stove.

Based on earlier studies (e.g., Pillarisetti et al. 2014; Ruiz-Mercado et al. 2011, 2013) and our feasibility-study, we anticipated that many households would use both their three-stone fire and their Envirofit. One motivation for this is that common cooking practices in the area require two simultaneous cooking pots (for example, for rice and beans, or for *matooke* and a sauce), and the Envirofit heats only one pot. We were interested in whether having a second fuel-efficient stove would substantially end stove stacking. Thus, approximately four weeks after late buyers received their first Envirofit, we surprised both groups with the gift of a second Envirofit stove.

In short, during the first study wave, both early and late buyers had only three-stone fires; in the second study wave, early buyers had one Envirofit, along with their three-stone fires, but late buyers only had three-stone fires; in the third study wave, both groups of buyers had one Envirofit; and in the fourth wave, both early buyers and late buyers had two Envirofits. See Table 1 for the steps of the experimental rollout. We tracked stove temperatures for approximately 18 weeks

(May–September 2012). Each household had as many as two three-stone fires and two Envirofit stoves monitored with SUMs. By the end of the study, numerous SUMs had been lost or burned up; therefore, after we delivered the second Envirofit stove, we encountered a shortage of SUMs, so we focused measurement on both Envirofits and the primary three-stone fire.

C. SUMs

We installed small, inexpensive, and unobtrusive SUMs to record stove temperatures.[8] Ruiz-Mercado et al. (2008) initially suggested using SUMs to log stove temperatures, and various studies have used that method (Mukhopadhyay et al. 2012; Pillarisetti et al. 2014; Ruiz-Mercado et al. 2013). We installed SUMs on two Envirofits and two three-stone fires in each household when possible (recall that by the end of the study, numerous SUMs had been lost or burned up; therefore, only a few secondary three-stone fires were measured when all users had two Envirofits).

Throughout the study, field staff recorded about 2,400 visual observations of whether a stove was in use (on/off) when they visited homes. Also, we examined the temperature data immediately before and after the 2,400 visual observations of stove use. After understanding how temperature patterns changed at times of observed stove use, we developed an algorithm to predict cooking behaviors for the wider dataset of 1.7 million temperature readings during which we did not have visual observations. By "cooking," we mean that the algorithm predicts

[8]The SUMs used for our project, iButtons™ manufactured by Maxim Integrated Products, Inc., are small stainless steel temperature sensors about the size of a small coin and the thickness of a watch battery. Our SUMs recorded temperatures up to 85°C with an accuracy of +/– 1.3°C. For additional details see: http://berkeleyair.com/services/stove-use-monitoring-system-sums/. The SUMs cost approximately US$16 each. They recorded a temperature data point every 30 minutes for 6 weeks in a household before needing minimal servicing from a technician to download the data and reset the device.

stove use, not necessarily that a cook is standing above the fire and actively working on a meal. Our algorithm would likely detect "cooking" in cases of banking hot coals for the next meal, and while this is not a formal act of cooking, it does burn wood and increase particulate matter in the kitchen. This process, detailed in Simons et al. (2014a), allowed us to unobtrusively and inexpensively track daily stove usage on a large sample of households throughout the study. Appendix A provides additional details on placing SUMs, the process of converting temperature readings into measures of predicted cooking, and documents that SUMs attrition was random.

D. Kitchen Performance Tests and Particulate Matter Monitoring

We performed standard kitchen performance tests (KPTs) (Bailis, Smith, and Edwards 2007) in each household to measure the quantity of fuel wood used, record detailed food diaries of what households cooked, and measure household air pollution before any Envirofits were distributed, that is, when early buyers had one Envirofit and when both groups of buyers had two Envirofits. The KPT lasted approximately 72 hours and involved daily visits by a small team of researchers who weighed wood and collected food diaries, which record cooking and stove usage over the previous 24 hours. Households were asked to only use wood from a specific pile so that the team could determine the change in weight over each day. In the food diary, households recorded what foods were cooked for each meal.

During household visits, we also monitored household air pollution. Residential combustion of solid fuels in developing countries is a significant source of pollutants that harms both the climate and health (Bond et al., 2004; Smith et al, 2004). Roughly 10%–38% of the carbon contained in fuels is not completely combusted when used in simple cooking technologies (Zhang et al., 2000). The carbon that is not converted into CO_2 is instead emitted as products of incomplete combustion (PICs) that contain potent health-damaging pollutants. We measured household level particulate matter (PM2.5) concentrations over the same 72 hours of the KPT. To measure PM2.5, we used the University of California, Berkeley (UCB) Particle and Temperature Sensor, which is a small, portable data logging device (a modified commercial smoke detector) that uses an optical scattering sensor to measure real-time PM2.5 concentrations.[9]

E. Long-Term Stove Usage

We revisited households approximately 3.5 years after they initially received their Envirofit stoves. The survey team made quick, unannounced, observation visits in November 2015 to see whether Envirofit stoves were still in use. The purpose of the visits was to observe which stoves were in use at the time of the visit, examine Envirofits and three-stone fire locations for obvious signs of use (smoke stains, black soot, etc.), and ask a series of short qualitative consumer satisfaction questions about the different stove types. We observed 82% (137 of 168) of the households.

III. Specification

We analyzed wood usage (kg/day), daily household air pollution (PM2.5) concentrations, and stove usage. Recall that there were four study waves with different levels of stove ownership: (1) households that had two three-stone fires; (2) early buyers who had received an Envirofit and late buyers who had only their three-stone fires; (3) both groups of buyers that had one Envirofit; (4) both groups

[9] The UCB Particle Monitor User Manual (Berkeley Air Monitoring Group and Indoor Air Pollution Team, School of Public Health, University of California 2010) details how to use these sensors.

of buyers that had received a second Envirofit. Due to budgetary constraints, we could only run KPTs at phases (1), (2), and (4). Thus, for outcomes measured in KPTs (wood usage, PM2.5), the regression specification using data from study waves (1), (2), and (4) was as follows:

$$(1)\ Y_{ipt} = \alpha_{ip} + b_0 * T_i + b_1 * Early\_have\_Envirofit_t + b_2 * Both\_have\_two\_Envirofits_t + \beta_1 (T_i * Early\_have\_Envirofit_t) + \beta_2 (T_i * Both\_have\_two\_Envirofits_t) + \epsilon_{ipt},$$

where Y_{ipt} is daily wood use or daily PM2.5 concentrations for household i for parish p in study wave t, α_{ip} are fixed effects for each household, $Early\_have\_Envirofit_t$ and $Both\_have\_two\_Envirofits_t$ are dummies for the study wave, and T_i is a dummy equal to one if, in the early treatment group, ϵ_{ipt} is a residual that is clustered by the parish * study wave but is assumed to be independent and identically distributed (i.i.d.) within a parish and study wave. The coefficients of interest are β_1 (the effect of being in the early buyer group during the study wave [2], or the effect of owning an Envirofit while the comparison group has only three-stone fires), and β_2 (the effect of being in the early buyer group during study wave [4], or the effect of owning your first Envirofit for approximately 4 weeks longer than the comparison group when both groups own two Envirofits).

We also ran this equation without household fixed effects, but our preferred specification included them. The household fixed effect controls for unobserved characteristics of the household, such as the talent and cooking style of the household cook, and structural features of the kitchen, such as windows or ventilation. Because particulate matter has extreme positive outliers, we analyzed the natural log of PM2.5 (as is typical in studies that examine PM2.5). We also top and bottom coded PM2.5 at the 2nd and 98th percentiles, and top coded wood usage at the 98th percentile.

For stove usage, we had data for both during and between the three weekly periods when we measured wood usage and PM2.5. Thus, the regression specification for the SUM usage data was:

(2) $Y_{ipt} = \alpha_{ip} + b_0 * T_i + b_1 * Early\_have\_Envirofit_t + b_2 * Both\_have\_Envirofit_t + b_3 * Both\_have\_two\_Envirofits_t + \beta_1 (T_i * Early\_have\_Envirofit_t) + \beta_2 (T_i * Both\_have\_Envirofit_t) + \beta_3 (T_i * Both\_have\_two\_Envirofits_t) + \epsilon_{ipt}$,

where Y_{ipt} is daily three-stone fire or Envirofit usage derived from SUM readings for household i for parish p in study wave t, α_{ip} are fixed effects for each household, $Early\_have\_Envirofit_t$, $Both\_have\_Envirofit_t$, and $Both\_have\_two\_Envirofits_t$ are dummies for the study wave, and T_i is a dummy equal to one if, in the early treatment group. ϵ_{ip} is a residual that may be clustered by the parish * study wave but is assumed to be i.i.d. within a parish and study wave. The coefficients of interest are β_1 (the effect of being in the early buyer group during study wave [2], or the effect of owning an Envirofit while the comparison group has only three-stone fires), β_2 (the effect of being in the early buyer group during study wave [3], or the effect of owning your first Envirofit for approximately 4 weeks longer than the comparison group which also owns one Envirofit), and β_3 (the effect of being in the early buyer group during study wave [4], or the effect of owning your first Envirofit for approximately 4 weeks longer than the comparison group when both groups own two Envirofits).

A. Accounting for the Hawthorne Effect

Wood usage and PM data are only from field technicians' visits during the approximately 72-hour KPT measurement week. In a companion paper (Simons et al. 2017), we found that there was a significant Hawthorne effect during those weeks.[10] In an attempt to account for this effect, we calculated differences in stove usage between weeks when observers were present and weeks when they were not present and adjusted wood and PM2.5 measures as follows.

Let the subscript *group* = early or late buyer, and let the superscript *wave* = the experimental wave (i.e., [1] households with two three-stone fires; [2] early buyers with an Envirofit and late buyers only with three-stone fires; [3] both groups of buyers with one Envirofit; and [4] both groups of buyers with two Envirofits). Our estimate of wood usage adjusted for the Hawthorne effect was:

$$(3)\ \Delta Adj\_Wood_{group}^{wave} = \Delta TSF\_Hours_{group}^{wave} * \left(\frac{TSF\_Wood}{hour}\right) + \Delta ENV\_Hours_{group}^{wave} * \left(\frac{ENV\_Wood}{hour}\right).$$

Δ*TSF_Hours* and Δ*ENV_Hours* are the differences in hours cooked due to the Hawthorne effect on the three-stone fire (Envirofit) among those that own Envirofits. *TSF_Wood* per hour is wood consumption from the first KPT (when no one had an Envirofit) divided by cooking on the three-stone fires during those days. We did not have any periods when households only had Envirofits. Thus, we used the laboratory results (Figure 2) indicating that *ENV_Wood* per hour is half that of a three-stone fire.

For PM concentrations, we followed the same technique, and the Hawthorne-adjusted PM2.5 generated for each group of buyers was:

[10]We compared stove usage in KPT weeks when observers were present with stove usage in adjacent weeks with no observers and found that participants increased usage of Envirofits by about 3.0 hours per day and decreased usage of the primary three-stone fires by about 1.8 hours per day during the endline KPT (when households owned two Envirofits), but then reverted to previous usage patterns once the observers left (Simons et al. 2017). Also, see Garland, Gould, and Pennise (2018) for an additional example of observer-induced behavioral differences in stove use during kitchen monitoring periods.

$$(4)\ \Delta Adj\_PM2.5_{group}^{wave} = \Delta TSF\_Hours_{group}^{wave} * \left(\frac{TSF\_PM2.5\_Generated}{hour}\right) + \Delta ENV\_Hours_{group}^{wave} * \left(\frac{ENV\_PM2.5\_Generated}{hour}\right).$$

TSF_PM2.5_Generated per hour is calculated by dividing PM2.5 concentrations by three-stone fire use from the first kitchen performance test (when no one had an Envirofit). *ENV_PM2.5_Generated* per hour is from laboratory results (Figure 2).

Because we had sensor-based usage metrics that covered all weeks of the experiment, the estimates for changes in cooking behaviors (hours cooked per day on three-stone fires and Envirofits) from Eq. (2) were not likely affected by the observer-induced behavioral response.[11] However, because technicians were in homes to measure wood usage and PM2.5, we adjusted for the Hawthorne effect by using Eqs. (3) and (4).

IV. Results

A. Summary Statistics and Randomization Tests

Table 2 shows baseline summary statistics and balance tests for covariates. Randomization between early buyers and late buyers was successful. Only one difference among the 20 covariates was (weakly) statistically significantly different than zero. Participants who randomly received their Envirofits early had a higher value of assets (US$1,158 vs. US$905) (p=0.08). Control households used approximately 9.3 kg of daily wood, had an average PM2.5 reading of 414.3 $\mu g/m^3$ in their kitchens, and cooked for about 6.2 people.

[11] Observers (technicians) were only present in households in three 72-hour periods over the 18 weeks that sensors measured stove usage.

Households used their first Envirofit about 4.3 hours per day and their second Envirofit about 2.9 hours per day (Table 3).

B. Effects of Envirofits on Fuel Use and Pollution

We began by analyzing the causal impact of the introduced Envirofit stove on wood usage (Table 4) during our experiment. In the pre-intervention period, the control group used about 9.3 kg of wood/day (Table 2, column 1); these usage rates fell when the early group had one Envirofit (-1.9 kg/day, $p<0.01$, Table 4, column 1) and when both groups had two Envirofits (-2.5 kg/day, $p<0.01$, Table 4, column 1), but there were no statistically significantly different rates of reduction for those that had received their Envirofit in the early group. In our preferred specification, with household fixed effects (column 2), the early receipt of an Envirofit was causally associated with a change of about -1.1 kg/day, ($p<0.1$). This reduction in wood consumption was a modest reduction of about 12% from the pre-intervention control group wood usage level. When all owned two Envirofits, both groups reduced their wood usage by about 2.5 kg/day ($p<0.01$) or 27%, relative to the pre-intervention control group, with no statistically significant difference between groups.

In Table 5, we present the causal effects of the introduction of Envirofit stoves on household air pollution concentrations. Pre-intervention, the control group had a daily concentration of PM of about 414 $\mu g/m^3$ (Table 2, column 1). In our preferred specification with household fixed effects (Table 5, column 2), the introduction of the first Envirofit reduced PM concentrations by 12% ($p<.01$) compared to the control group. When both groups had two Envirofits, both groups reduced PM by about 10% ($p<0.1$) with no difference between groups. That is, having the first Envirofit longer did not result in detectably different pollution levels once both groups had received two Envirofits.

C. Effects of Envirofits on Cooking Behaviors

Next, we examined the effects of the introduction of Envirofits on daily time spent cooking on the existing three-stone fires. We had stove usage data for much longer periods than the three kitchen measurement periods. We estimated how the daily hours cooked on each stove varied over the entire 18 weeks of the study period (Table 6, based on Eq. 2). Figure 3 summarizes stove usage by study phase. A weekly time series of stove usage is shown Figures 4 and 5.[12]

Total usage on both three-stone fires was 12.7 hours per day by the control group in the sample of all weeks prior to Envirofit introduction. In our preferred specification (Table 6, column 2), the causal estimates were that the introduction of the first Envirofit reduced cooking on three-stone fires by about 3.7 hours per day ($p<0.01$). This was a reduction of about 30% from the control group prior to the introduction of the first Envirofit.

When late buyers received their first Envirofit (Table 6, column 2), we saw a reduction in use of the three-stone fires among late buyers by 3.1 hours per day ($p<0.01$) (about 25%); however, at the same time, we saw an increase in three-stone fire use of about 2.9 hours per day ($p<0.01$) (about 23%) in the early buyers (who had owned their Envirofits about 4 weeks longer than the late buyers). It is unclear why these differed in direction, though one possibility is that, after initial experimentation with the Envirofit, the early group had decided to use their three-stone fires more, while the late group continued to experiment with the new Envirofit. This difference appears to have resolved itself once both groups received their second Envirofit (Table 6, column 2), as combined use of the three-stone fires declined by about 5.2 hours per day ($p<0.01$, with no statistically

[12] See Appendix Figures A1 and A2 for the daily time series of stove use by early and late buyers, respectively.

significant difference if households received their first Envirofit earlier or later). This was a reduction of about 41% in three-stone fire use once both Envirofits were introduced. In short, even with two Envirofit stoves, most households continued to use their three-stone fires regularly.

D. Adjusting for the Hawthorne Effect

To adjust for this effect, we calculated the change in three-stone fire and Envirofit hours cooked in the measurement week compared to all weeks.[13] To do this for three-stone fires, we ran the regression for the effect of the Envirofit on hours cooked on three-stone fires, but restricted the sample to only observations during the measurement week (Table 7). The difference of the coefficients between Table 6 (all weeks) and Table 7 (only measurement weeks) was the delta three-stone fire hours used in Eqs. (3) and (4). To calculate the change in hours cooked on Envirofits, we ran similar regressions, but instead used hours cooked on the Envirofit as the dependent variable (Table 8 [all weeks] and Table 9 [measurement weeks]).

Use of three-stone fires fell by 6.4 hours per day when the first Envirofit was delivered, when only looking at the week when the KPTs were performed (Table 7, column 2), versus 3.7 hours per day over the entire period with sensors (Table 6, column 2). Usage of the Envirofit was roughly 3.8 hours per day when the first Envirofit was delivered, when only looking at the kitchen measurement week (Table 9, column 3), versus 1.5 hours per day over the entire period with sensors (Table 8, column 3). This reduction in three-stone fire use and increase in

[13]Note that this is one option for addressing the Hawthorne effect. As this is not a methodological paper, we only show this option, but we realize that other options are reasonable (e.g., only use one week before/after observers are present to adjust estimated use). Thus, we add the caveat that this method is only a rough estimation of the Hawthorne effect on differences in wood use and particulate matter.

Envirofit use was anticipated because the measurement weeks had the Hawthorne effect resulting from the daily visits of our enumerators (Simons et al. 2017).

Thus, we adjusted for the 2.6 hours per day increased use of three-stone fires and 2.4 hours per day (Table 10) decreased use of one Envirofit outside of the measurement week using Eqs. (3) and (4). This adjustment yielded a smaller estimated reduction in wood use: 1.7% (Table 11, first panel) as opposed to the unadjusted reduction of 11.6% (Table, 4, column 2). We also found a smaller reduction of PM2.5: 0.3% (Table 11, second panel) instead of the unadjusted reduction of 12.0% (Table 5, column 2).

Next, we calculated the Hawthorne adjustment for the periods when participants had two Envirofits. Use of three-stone fires fell 10.2 hours per day when participants had two Envirofits during the measurement week (Table 7, column 2), versus 5.2 hours per day during the entire period with sensors (Table 6, column 2). Use of the Envirofits was 6.8 hours per day during the measurement week (Table 9, columns 3 and 5), versus 3.7 hours per day during the entire period with sensors (Table 8, columns 3 and 5).[14] Therefore, we adjusted for the 5.1 hours per day increased use of the three-stone fires and 3.1 hours per day (Table 10) decreased use of two Envirofits outside of the measurement week using Eqs. (3) and (4). The estimate of daily wood use changed from an unadjusted reduction of 26.7% to a reduction of 2.5% after the adjustment (Table 11, panel one). The estimate of daily PM2.5 concentrations changed from an unadjusted reduction of 10.0% to an increase of 18.3% after the adjustment (Table 11, panel two).

[14]We calculated total Envirofit cooking as the sum of cooking on the first Envirofit plus the cooking on the second Envirofit individually, because only about 60% of the households had any combined readings from both SUM devices during the final measurement week.

E. Long-term Usage

We made unannounced visits to measure stove usage approximately 3.5 years after the initial Envirofit stoves were distributed. Approximately 82% of the original households were home when we visited.

At the exact moment our enumerators arrived, about 48% (66 out of 137) of the households were actively cooking (Table 10). Among those, only 9% (6 out of 66) were cooking with an Envirofit stove. Enumerators asked the 131 households that were not cooking on the Envirofit when enumerators arrived if they could inspect their Envirofit to see obvious signs of use, such as black soot or fresh ashes in the stove (Figure 6 shows an example of a stove with obvious signs of use). Among those households, 65% had an Envirofit with obvious signs of use, 17% had Envirofits stored that were clearly not being used, 2% had Envirofits that were still in perfect condition (essentially never used), 8% said their Envirofit was damaged and disposed of, and a final 8% said they had given the stove away. Next, enumerators asked households to see their second Envirofit to determine if it had signs of use. Among this sample, 25% had a second Envirofit with obvious signs of use, 11% had their second Envirofit stored with limited signs of use, 9% had a second Envirofit that had never been used, 38% reported they had given the second Envirofit away as a gift, and 16% said the second Envirofit was damaged and they disposed of it.

Among all households visited (N=137), 23% reported that they still used both Envirofits, 50% said they used only one Envirofit, and 27% said they had stopped using Envirofits completely.

Enumerators also asked all households if they had to purchase a stove today, would they purchase an Envirofit. Among respondents, 79% said they would purchase an Envirofit, and 15% said they would not purchase an Envirofit, with the remaining households unsure. Given that the share that stated a willingness to repurchase was greater than the share using the Envirofit, we suspected this self-report was biased.

Enumerators then asked open-ended response questions as to the reasons for those hypothetical purchase decisions. The most popular responses among those that would buy another Envirofit were that the stove saved fuel and reduced household time collecting fuel, the stove cooked fast, the stove was easily portable, and the stove produced less smoke than a three-stone fire. Among those that said they would not purchase another Envirofit, the most popular responses were that the preparation of firewood was difficult for Envirofits (needed smaller pieces of wood than a three-stone fire), the stove did not simmer food, the stove was too small for the household's cooking needs, it was hard to prepare some traditional meals on the stove, and the stove was hard to light.

F. Rebound Effects

Rebound effects occur when improvements in energy efficiency make consuming energy less expensive and therefore encourage increased consumption of energy (see review in Sorrell, Dimitropoulos, and Sommerville [2009]). While we did not have fuel cost data to formally estimate a rebound effect, we examined stove use graphically, as shown in Figure 3, which suggested the presence of a rebound effect. When households first received an Envirofit, they reduced three-stone fire usage. However, by the end of our tracking period, Envirofit usage had increased more than three-stone fire use had decreased. The aggregate time all stoves were in use increased by about 20% throughout the period that we tracked stove temperatures.

V. Discussion and Conclusion

This study was the first randomized trial that collected detailed stove usage metrics among households that paid market prices for their stoves. We found a slight reduction in wood use (-11.6%) and PM2.5 concentrations (-12.0%) after the introduction of one Envirofit, but this reduction mostly vanished if we adjusted for the Hawthorne effect.

Despite our selection of a sample that paid market price for their fuel-efficient stove, it did not appear that usage rates of the new stove were markedly different than studies that offered highly subsidized stoves. For example, in Pillarisetti et al. (2014), which also used temperature sensors to track detailed stove level usage, households received fuel-efficient stoves for free and ended up using their traditional stoves about 75% of the time and the introduced fuel-efficient stove about 25% of the time. Our results were very similar, with roughly 67% of cooking done on the three-stone fires and 33% on fuel-efficient surfaces by the end of our study. Hanna, Duflo, and Greenstone (2016) did not gather stove use monitor data; however, their conclusion was that fuel-efficient stove use was enough to reduce indoor air pollution in the initial phase of their experiment, but that in the longer term, poor maintenance of the stoves led to an elimination of the air pollution benefits. Our results were similar, except that, in our follow-up, it did not appear that a lack of stove durability was the cause of limited stove use.

A second innovation in our study was to see if households would fully switch from the traditional smoky cookstove, if given a second Envirofit. Despite the second fuel-efficient cooking surface, households continued to mostly use the traditional cookstove. Almost all households used both three-stone fires and fuel-efficient stoves in daily cooking. It appeared that households used the fuel-efficient stove to heat things that cook relatively quickly, such as boiling water to make tea and sauces. They preferred three-stone fires for low-heat cooking, such as simmering dishes like beans and cooking bananas. It appeared that the ability to modulate the stove's temperature would be a valued feature for cooks.

Our third contribution was measuring the bias caused by observer-induced bias, or the Hawthorne effect. By collecting stove temperature data when technicians were in the home and comparing it to times they were not in the home, we found that households cooked about 2.5 hours per day more on the Envirofit and 2.5 hours less per day on three-stone fires when observers were present and then switched back to previous patterns once the observers had left. We found reductions in wood use (-11.6%) and PM2.5 concentrations (-12.0%) after the introduction of one Envirofit, but once we adjusted for the different behavior when observers were present, this reduction was almost zero. In regard to impacting environmental and health problems, fuel use and particulate matter would need to have declined by much more than what was found in this study. To reach WHO targets for household air pollution, particulate matter needed to decline by 90% from pre-intervention levels. Throughout the study period, three-stone fire use fell by about 2.5 hours a day, but this was more than offset by about 5 hours a day of new cooking on the introduced stoves. This increase in total cooking time diminishes the environmental and household air pollution benefits compared to those shown in the laboratory results. While any reduction in fuel use and particulate matter was likely beneficial for households,[15] fuel-efficient wood stoves such as these will not be adequate to reach safe levels of household air pollution. Thus, policies that assist consumers to shift to safe fuels such as gas or electricity—particularly when coupled with policies to disable smoky indoor stoves—should take on increased importance.

[15]Emerging evidence shows that small reductions in PM2.5 can have benefits in especially vulnerable subpopulations. For example, even a small reduction in PM2.5 can reduce adverse pregnancy outcomes (Alexander et al. 2018) and improve growth in children under the age of two years (LaFave et al. 2019).

APPENDIX A

The details presented here summarize our previous research on how we converted temperature readings into stove usage metrics and measured if the attrition of stove use monitors was random (this appendix is based on Harrell et al. 2016; Simons et al. 2014; 2017; 2018).

A. Placement of SUMs

SUMs must be close enough to the heat source to capture changes in temperatures, but not so close that they exceed 85°C, the maximum temperature the SUMs used in this study can record before they overheat and malfunction. We do not need to recover the exact temperature of the hottest part of the fire to learn about cooking behaviors. Even with SUMs that are reading temperatures 20–30 cm from the center of the fire, as long as the temperature readings for times when stoves are in use are largely different than times when stoves are not used, the logistic regression will be able to predict a probability of usage.

SUMs for three-stone fires were placed in a SUM holder (Figure A3) and then placed under one of the stones in the three-stone fire (left panel, Figure A4). The SUMs for Envirofits were attached using duct tape and wire and placed at the base of the stove behind the intake location for the firewood (right panel, Figure A4). Figure A5 shows an example of SUM temperature data for a household over about three weeks. The left panel shows the temperatures registered in a three-stone fire versus the ambient temperature also recorded with SUMs in this household, while the right panel compares the temperature of the Envirofit to the ambient temperature reading.

B. Visual Observations of Use

Each time data collection personnel visited a household; they observed which stoves were in use (whether the stove was "on" or "off," along with the date and timestamp recorded digitally via a handheld device). Enumerators visited each household several times during a "measurement week," when they also enumerated a survey and weighed wood for the KPT. Another enumerator visited once every 4 to 6 weeks to download data and reset the SUM devices.

C. Generating an Algorithm

We matched the observations of stove use to SUM temperature data by time- and date stamps. At the core of our method was a logistic regression using the lags and leads of the SUM temperature data to predict visual observations of stove usage. We tested 10 specifications of differing combinations of current, lagged, and leading temperature readings (Simons et al. 2014).

In order to determine which of the models was most appropriate, we tested the 10 specifications with the Akaike information criterion (AIC) (Akaike 1981). The AIC trades off goodness of fit of the model with the complexity of the model to guard against over-fitting.

The preferred specification included the temperature reading closest to the time of the observation, the readings 60 and 30 minutes prior, and 60 and 30 minutes after the observation of use, and a control for hour of the day. This regression specification correctly predicted 89.3% of three-stone fire observations and 93.8% of Envirofit observations of stove usage. We then compared our algorithm to other previously published algorithms (Mukhopadhyay et al. 2012; Ruiz-Mercado, Canuz, and Smith 2012). Those algorithms focused on defining "discrete" cooking events based on rapid temperature slope increases and elevated stove temperatures, followed by a cooling off period. We applied those algorithms to the temperature data we had collected and found our logistic regression correctly classified more observations, with a higher pseudo R-squared, than any other algorithm for both three-stone fires and the Envirofits.

D. Random SUMs Attrition

One concern for our study is whether the attrition of the sensors used to measure stove temperatures was random. In cases of sensor malfunction we lost the temperature readings associated with that device (about six weeks of data for that individual stove). The concern is that if damage (overheating above the 85°C tolerance of our SUMs device) was more likely on stoves that were used more heavily, then the data we have are not an unbiased measure of stove usage for the broader sample. If however, the attrition of SUMs sensors is random, there is less concern about the internal validity of our sample.

To examine this topic we follow the approach outlined in Simons et al. (2017) and focus on the endline period where all participants had two three stone fires and two Envirofit stoves. We test this in various ways. First, we regress whether the SUMs data was missing at endline (device malfunctioned) on household fuel wood consumption during that same period. Because fuel wood is a direct input into how much the stoves were used, this is the most direct test of this relationship. If households that cook more (using fuel wood consumption as a proxy) also have a higher probability of SUMs attrition, this would be evidence of non-random attrition and a problem for our study. We examine this relationship separately for each stove type that we included in our study (recall that we choose not to track the non-primary three stone fire by the endline of our study).

Because we are testing for attrition due to excessive cooking (heat exposure) we only test for this relationship on the sample of stoves on which we placed a SUMs device. We also do similar checks with other variables that are related to cooking or experience (count of people cooked for daily, number of meals cooked daily, number of meals in which *matooke* was cooked daily, and age of the cook).

In Table A1 we present the results of the attrition checks. In our preferred test, we find that the likelihood of SUMs survival is statistically no different than zero (col. 1-3) for each additional kilogram of wood consumption. When examining whether a larger household size is associated with the likelihood of SUMs survival we find a weakly statistically significant relationship for primary three stone fire usage (col. 4). Each additional person cooked for is associated with a four-percentage point decrease in the probability of SUMs survival ($p<0.10$), however this relationship does not appear for either of the Envirofit stoves (col. 5-6). Lastly, we test whether the count of daily meals cooked (col. 7-9), daily meals in which *matooke* was cooked (col. 10-12), or the age of the cook (col. 13-15) is associated with SUMs survival. We find no statistically significant relationship. Taken as a whole, these tests do not provide strong evidence of non-random attrition of SUMs devices.

Acknowledgements: The study was funded by the United States Agency for International Development under Translating Research into Action, Cooperative Agreement No. GHS-A-00-09-00015-00. The grant recipient was Impact Carbon, which co-funded and managed the project. This study received Institutional Review Board approval from the Committee for Protection of Human Subjects (approval number 2010-06-1665) at the University of California, Berkeley and Cornell University. This study is registered in the AEA RCT Registry, and the unique identifying number is: AEARCTR-0003778. We would like to thank the Atkinson Center for a Sustainable Future at Cornell University, the Institute for the Social Sciences at Cornell University, the Cornell Population Center, and the National Science Foundation INFEWS program for additional funding of related expenses. Data collection was carried out by the Center for Integrated Research and Community Development (CIRCODU), and the project's success depended greatly on its managers—Joseph Arinieitwe Ndemere, Juliet Kyaesimira, Vastinah Kemigisha—and field supervisors—Innocent Byaruhanga, Fred Isabirye, Michael Mukembo, Moreen Akankunda, and Noah Kirabo. We thank Impact Carbon partners Matt Evans, Evan Haigler, Jimmy Tran, Caitlyn Toombs, and Johanna Young; U.C. Berkeley Household Energy, Climate, and Health Research Group partners including Kirk Smith, Ilse Ruiz-Mercado, and Ajay Pillarisetti; Berkeley Air partners including Dana Charron, David Pennise, Michael Johnson, and Erin Milner; and the USAID TRAction Technical Advisory Group. The views expressed in this paper solely reflect those of the authors, and these opinions are not necessarily those of the institutions with which the authors are affiliated. Declarations of interest: none. All errors are our own.

REFERENCES

Akaike, Hirotugu. 1981. "Likelihood of a Model and Information Criteria." *Journal of Econometrics* 16: 3–14.

Alexander, Donee A., Amanda Northcross, Theodore Karrison, Oludare Morhasson-Bello, Nathaniel Wilson, Omolola M. Atalabi, Anindita Dutta, et al. 2018. "Pregnancy Outcomes and Ethanol Cook Stove Intervention: A Randomized-Controlled Trial in Ibadan, Nigeria." *Environment International* 111: 152–163. https://doi.org/10.1016/j.envint.2017.11.021.

Ashraf, Nava, James Berry, and Jesse M Shapiro. 2010. "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia." *American Economic Review* 100 (5): 2383–2413. https://doi.org/10.1257/aer.100.5.2383.

Bailis, Rob, Kirk R. Smith, and Rufus Edwards. 2007. "Kitchen Performance Test (KPT)." Univeirsity of California, Berkeley, CA.

Bailis, Robert, Rudi Drigo, Adrian Ghilardi, and Omar Masera. 2015. "The Carbon Footprint of Traditional Woodfuels." *Nature Climate Change* 5 (3): 266–272. https://doi.org/10.1038/nclimate2491.

Beltramo, Theresa, Garrick Blalock, David I. Levine, and Andrew M. Simons. 2015a. "Does Peer Use Influence Adoption of Efficient Cookstoves? Evidence From a Randomized Controlled Trial in Uganda." *Journal of*

Health Communication 20 (S1): 55–66. https://doi.org/10.1080/10810730.2014.994244.
———. 2015b. "The Effect of Marketing Messages and Payment over Time on Willingness to Pay for Fuel-Efficient Cookstoves." *Journal of Economic Behavior & Organization* 118: 333–45. https://doi.org/10.1016/j.jebo.2015.04.025.
Bensch, Gunther, and Jörg Peters. 2015. "The Intensive Margin of Technology Adoption – Experimental Evidence on Improved Cooking Stoves in Rural Senegal." *Journal of Health Economics* 42: 44–63. https://doi.org/10.1016/j.jhealeco.2015.03.006.
Berkeley Air Monitoring Group, and Berkeley Indoor Air Pollution Team, School of Public Health, University of California. 2010. "UCB Particle Monitor User Manual." Berkeley, CA. http://edge.rit.edu/edge/P13625/public/Reference%20Documents/User%20Manual_UCB%20Particle%20Monitor_v8%2030Dec2010.pdf.
Bonjour, Sophie, Heather Adair-Rohani, Jennyfer Wolf, Nigel G. Bruce, Sumi Mehta, Annette Prüss-Ustün, Maureen Lahiff, Eva A. Rehfuess, Vinod Mishra, and Kirk R. Smith. 2013. "Solid Fuel Use for Household Cooking: Country and Regional Estimates for 1980-2010." *Environmental Health Perspectives* 121 (7): 784–790. https://doi.org/10.1289/ehp.1205987.
Chay, Kenneth Y., and Michael Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence From Geographic Variation in Pollution Shocks Induced by a Recession." *Quarterly Journal of Economics* 118 (3): 1121–1167. https://doi.org/10.1162/00335530360698513.
Chowdhury, Zohir, Rufus Edwards, Michael Johnson, Kyra Naumoff Shields, Tracy Allen, Eduardo Canuz, and Kirk R Smith. 2007. "An Inexpensive Light-Scattering Particle Monitor: Field Validation." *Journal of Environmental Monitoring* 9 (10): 1099–1106.
Cohen, Jessica, and Pascaline Dupas. 2010. "Free Distribution or Cost-Sharing? Evidence from a Malaria Prevention Experiment." *Quarterly Journal of Economics* 125 (1): 1–45.
Currie, J, and WR Walker. 2009. "Traffic Congestion and Infant Health: Evidence from E-Zpass." *American Economic Journal: Applied Economics* 3 (1): 65–90. http://www.nber.org/papers/w15413.
Dickinson, Katherine L, Ernest Kanyomse, Ricardo Piedrahita, Evan Coffey, Isaac J Rivera, James Adoctor, Rex Alirigia, et al. 2015. "Research on Emissions, Air Quality, Climate, and Cooking Technologies in Northern Ghana (Reaccting): Study Rationale and Protocol." *BMC Public Health* 15 (1): 126. https://doi.org/10.1186/s12889-015-1414-1.

Dupas, Pascaline. 2014. "Getting Essential Health Products to Their End Users: Subsidize, but How Much?" *Science* 345 (6202): 1279–81. https://doi.org/10.1126/science.1256973.

Ezzati, Majid, and Daniel M Kammen. 2001. "Indoor Air Pollution from Biomass Combustion and Acute Respiratory Infections in Kenya: An Exposure-Response Study." *Lancet* 358 (9282): 619–24. http://www.ncbi.nlm.nih.gov/pubmed/11530148.

———. 2002. "Evaluating the Health Benefits of Transitions in Household Energy Technologies in Kenya." *Energy Policy* 30 (10): 815–26. https://doi.org/10.1016/S0301-4215(01)00125-2.

Ezzati, Majid, Homayoun Saleh, and Daniel M Kammen. 2000. "The Contributions of Emissions and Spatial Microenvironments to Exposure to Indoor Air Pollution from Biomass Combustion in Kenya." *Environmental Health Perspectives* 108 (9): 833–39.

Fischer, Greg, Dean Karlan, Margaret McConnell, and Pia Raffler. 2019. "Short-Term Subsidies and Seller Type: A Health Products Experiment in Uganda." *Journal of Development Economics* 137: 110–24. https://doi.org/doi.org/10.1016/j.jdeveco.2018.07.013.

Garland, C., C. F. Gould, and D. Pennise. 2018. "Usage and Impacts of the Envirofit HM-5000 Cookstove." *Indoor Air* 28 (4): 640–650. https://doi.org/10.1111/ina.12460.

Hanna, Rema, Esther Duflo, and Michael Greenstone. 2016. "Up in Smoke: The Influence of Household Behavior on the Long-Run Impact of Improved Cooking Stoves." *American Economic Journal: Economic Policy* 8 (1): 80–114.

Harrell, Stephen, Theresa Beltramo, Garrick Blalock, Juliet Kyayesimira, David I. Levine, and Andrew M. Simons. 2016. "What Is a Meal?: Comparing Methods of Auditing Carbon Offset Compliance for Fuel Efficient Cookstoves." *Ecological Economics* 128: 8–16.

LaFave, Daniel, Abebe Damte Beyene, Randall Bluffstone, Sahan T.M. Dissanayake, Zenebe Gebreegziabher, Alemu Mekonnen, and Michael Toman. 2019. "Impacts of Improved Biomass Cookstoves on Child and Adult Health: Experimental Evidence from Rural Ethiopia." Waterville, ME. http://web.colby.edu/drlafave/lafaveetal_stovehealth/.

Levine, David I, Theresa Beltramo, Garrick Blalock, Carolyn Cotterman, and Andrew M. Simons. 2018. "What Impedes Efficient Adoption of Products? Evidence from Randomized Sales Offers for Fuel-Efficient Cookstoves in Uganda." *Journal of the European Economic Association* 16 (6): 1850–80.

Lewis, Jessica J, and Subhrendu K Pattanayak. 2012. "Who Adopts Improved Fuels and Cookstoves? A Systematic Review." *Environmental Health Perspectives* 120 (5): 637–45. https://doi.org/10.1289/ehp.1104194.

Lim, Stephen S, Theo Vos, Abraham D Flaxman, Goodarz Danaei, Kenji Shibuya, Heather Adair-Rohani, Markus Amann, et al. 2012. "A Comparative Risk Assessment of Burden of Disease and Injury Attributable to 67 Risk Factors and Risk Factor Clusters in 21 Regions, 1990-2010: A Systematic Analysis for the Global Burden of Disease Study 2010." *Lancet* 380 (9859): 2224–60. http://www.ncbi.nlm.nih.gov/pubmed/23245609.

Masera, Omar R, Barbara D Saatkamp, and Daniel M Kammen. 2000. "From Linear Fuel Switching to Multiple Cooking Strategies: A Critique and Alternative to the Energy Ladder Model." *World Development* 28 (12): 2083–2103. https://doi.org/10.1016/S0305-750X(00)00076-0.

Mortimer, Kevin, Chifundo B Ndamala, Andrew W Naunje, Jullita Malava, Cynthia Katundu, William Weston, Deborah Havens, et al. 2017. "A Cleaner Burning Biomass-Fuelled Cookstove Intervention to Prevent Pneumonia in Children under 5 Years Old in Rural Malawi (the Cooking and Pneumonia Study): A Cluster Randomised Controlled Trial." *The Lancet* 389 (10065): 167–75. https://doi.org/10.1016/S0140-6736(16)32507-7.

Mukhopadhyay, Rupak, Sankar Sambandam, Ajay Pillarisetti, Darby Jack, Krishnendu Mukhopadhyay, Kalpana Balakrishnan, Mayur Vaswani, et al. 2012. "Cooking Practices, Air Quality, and the Acceptability of Advanced Cookstoves in Haryana, India: An Exploratory Study to Inform Large-Scale Interventions." *Global Health Action* 5: 19016–19016.

Pillarisetti, Ajay, Mayur Vaswani, Darby Jack, Kalpana Balakrishnan, Michael N Bates, Narendra K Arora, and Kirk R Smith. 2014. "Patterns of Stove Usage after Introduction of an Advanced Cookstove: The Long-Term Application of Household Sensors." *Environmental Science & Technology* 48 (24): 14525–14533.

Pope III, C. Arden, Richard T. Burnett, Michael J. Thun, Eugenia E. Calle, Daniel Krewski, and George D. Thurston. 2002. "Lung Cancer, Cardiopulmonary Mortality, and Long-Term Exposure to Fine Particulate Air Pollution." *Journal of the American Medical Association* 287 (9): 1132–1141. https://doi.org/10.1001/jama.287.9.1132.

Ramanathan, Veerabhadran, and Gregory Carmichael. 2008. "Global and Regional Climate Changes Due to Black Carbon." *Nature Geoscience* 1 (4): 221–227. http://www.nature.com/ngeo/journal/vaop/ncurrent/full/ngeo156.html.

Rehfuess, Eva A., David J. Briggs, Mike Joffe, and Nicky Best. 2010. "Bayesian Modelling of Household Solid Fuel Use: Insights towards Designing Effective Interventions to Promote Fuel Switching in Africa." *Environmental Research* 110 (7): 725–732. https://doi.org/10.1016/j.envres.2010.07.006.

Ruiz-Mercado, Ilse, Eduardo Canuz, and Kirk R. Smith. 2012. "Temperature Dataloggers as Stove Use Monitors (SUMs): Field Methods and Signal Analysis." *Biomass and Bioenergy* 47: 459–68. https://doi.org/10.1016/j.biombioe.2012.09.003.

Ruiz-Mercado, Ilse, Eduardo Canuz, Joan L. Walker, and Kirk R. Smith. 2013. "Quantitative Metrics of Stove Adoption Using Stove Use Monitors (SUMs)." *Biomass and Bioenergy* 57 (October): 136–48. http://dx.doi.org/10.1016/j.biombioe.2013.07.002.

Ruiz-Mercado, Ilse, Nick L Lam, Eduardo Canuz, Gilberto Davila, and Kirk R Smith. 2008. "Low-Cost Temperature Loggers as Stove Use Monitors (SUMs)." *Boiling Point* 55: 16–18.

Ruiz-Mercado, Ilse, Omar Masera, Hilda Zamora, and Kirk R. Smith. 2011. "Adoption and Sustained Use of Improved Cookstoves." *Energy Policy* 39 (12): 7557–66. https://doi.org/10.1016/j.enpol.2011.03.028.

Simons, Andrew M., Theresa Beltramo, Garrick Blalock, and David I. Levine. 2014. "Comparing Methods for Signal Analysis of Temperature Readings from Stove Use Monitors." *Biomass and Bioenergy* 70: 476–488.

———. 2017. "Using Unobtrusive Sensors to Measure and Minimize Hawthorne Effects: Evidence from Cookstoves." *Journal of Environmental Economics and Management* 86: 68–80. https://doi.org/10.1016/j.jeem.2017.05.007.

———. 2018. "Sensor Data to Measure Hawthorne Effects in Cookstove Evaluation." *Data in Brief* 18: 1334–1339. https://doi.org/10.1016/j.dib.2018.04.021.

Smith, Kirk R, Nigel Bruce, Byron Arana, Alisa Jenny, Asheena Khalakdina, John McCracken, Anaite Diaz, Morten Schei, and Sandy Gove. 2006. "Conducting The First Randomized Control Trial of Acute Lower Respiratory Infections and Indoor Air Pollution: Description of Process and Methods." *Epidemiology* 17 (6): S44–S44.

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." *Lancet* 378 (9804): 1717–26. https://doi.org/10.1016/S0140-6736(11)60921-5.

Smith-Sivertsen, Tone, Esperanza Díaz, Dan Pope, Rolv T Lie, Anaite Díaz, John 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: The RESPIRE Randomized Trial, Guatemala." *American Journal of Epidemiology* 170 (2): 211–20. https://doi.org/10.1093/aje/kwp100.

Sorrell, Steve, John Dimitropoulos, and Matt Sommerville. 2009. "Empirical Estimates of the Direct Rebound Effect: A Review." *Energy Policy* 37 (4): 1356–71. https://doi.org/10.1016/j.enpol.2008.11.026.

World Health Organization. 2006. "WHO Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide: Global Update 2005: Summary of Risk Assessment," 1–22. https://doi.org/10.1016/0004-6981(88)90109-6.

Figure 1
Comparison of wood burning stoves: three stone fire versus Envirofit G-3300

(a) Three Stone Fire

(b) Envirofit G-3300

Figure 2

Certified Emissions and Performance Report for Envirofit G3300

April 27, 2011

ENGINES & ENERGY CONVERSION LAB

DEPARTMENT OF
MECHANICAL ENGINEERING
COLORADO STATE UNIVERSITY

1374 CAMPUS DELIVERY
FORT COLLINS, CO
80523-1374
970.491.4796
970.491.4799 (F)
WWW.EECL.COLOSTATE.EDU

Emissions and Performance Report

The stove listed below has been tested in accordance with the "*Emissions and Performance Test Protocol*", with emissions measurements based on the biomass stove testing protocol developed by Colorado State University (available at www.eecl.colostate.edu). Percent improvements are calculated from three-stone fire performance data collected at Colorado State University.

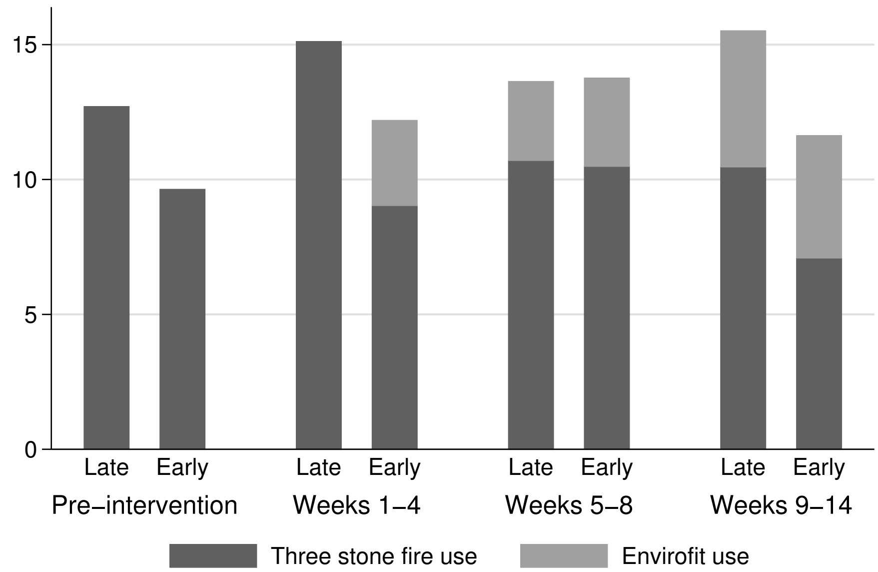
Stove Manufacturer:	Envirofit International
Stove Model:	G-3300
Test Dates:	4/4/2011-4/22/2011
Average CO emissions (grams):	18.7
80% Confidence Interval:	17.7-19.7
Percent Improvement:	65.30%
Average PM emissions (milligrams):	995
80% Confidence Interval:	944-1046
Percent Improvement:	51.20%
Average Fuel use (grams):	596.7
80% Confidence Interval:	591.6-601.7
Percent Improvement:	50.10%
Average Thermal efficiency:	32.6
80% Confidence Interval:	32.3-32.8
Percent Improvement:	105.20%
High Power (kW):	3.3
80% Confidence Interval:	3.3-3.4
Low Power (kW):	1.9
______ 80% Confidence Interval:	1.8-1.9

The above results are certified by the Engines and Energy Conversion Laboratory at Colorado State University. All claims beyond the above data are the responsibility of the manufacturer.

Morgan DeFoort
EECL Co-Director
Technical Lead, Biomass Stoves Testing Program

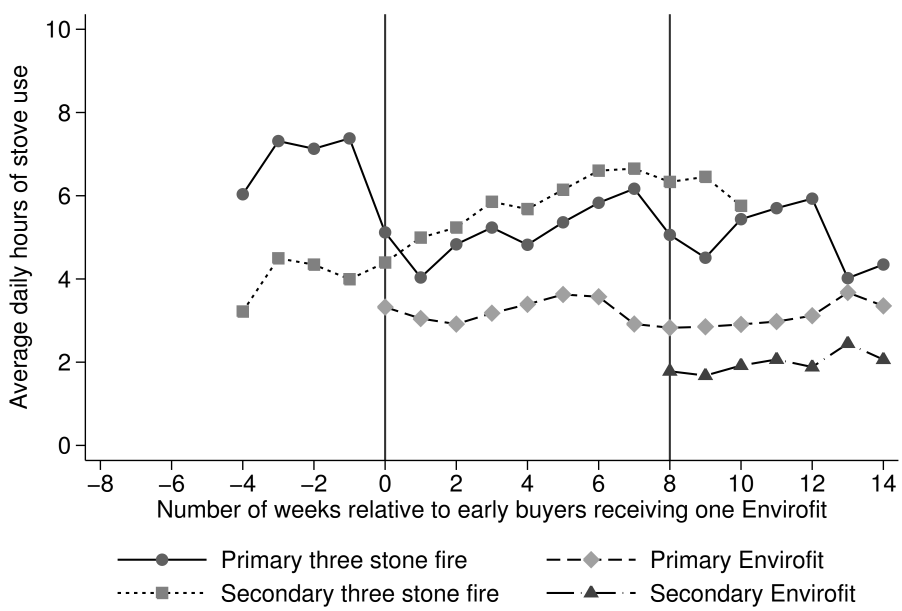
Note: The report can be downloaded at http://www.envirofit.org/images/products/pdf/g3300/G3300Cert.pdf

Figure 3
Average Daily Stove Use



Note: Pre−intervention (4 weeks) no Envirofits; Weeks 1−4 early buyers have one Envirofit; Weeks 5−8 all have one Envirofit; Weeks 9−14 all have two Envirofits.

Figure 4
Weekly Stove Use of Early Buyers



Note: Vertical lines designate when households received their first and second Envirofits.

Figure 5
Weekly Stove Use of Late Buyers

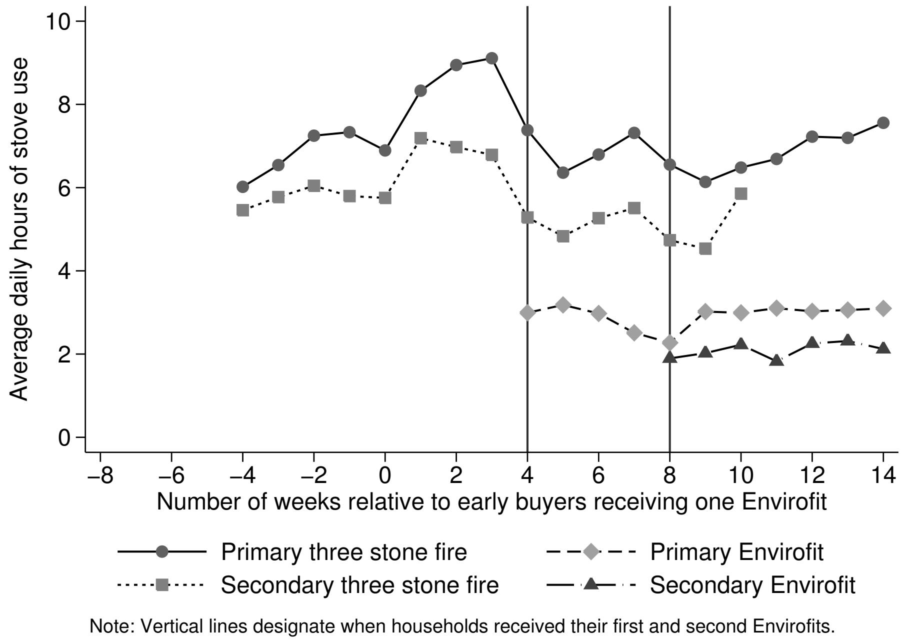


Figure 6
Envirofit Stove with Obvious Signs of Use (from Long Term Usage Study in Nov. 2015)

Table 1
Timeline of Experimental Rollout

Approximate Timing	**Event**
Weeks -4 to week 0	Stove use monitoring (SUMs) begins on two three stone fires
Week 0	Baseline kitchen performance tests (wood weighting) and particulate matter (PM2.5) monitoring*
End of week 0	Deliver first Envirofit to early buyers
Weeks 1-4	SUMs monitoring continues
Week 4	Midline kitchen perfomrance test and PM2.5 monitoring*
End of week 4	Deliver first Envirofit to late buyers (now all participants have one Envirofit)
Weeks 4-8	SUMs monitoring continues
Week 8	Deliver second Envirofit to both early and late buyers
Weeks 8-14	SUMS monitoring continues**
Week 14	Endline kitchen performance test and PM2.5 monitoring*
3.5 years later	Long-term usage follow up

Note: Measurement dates and timing are approximate as roll-out was staggered across the 14 parishes. Stove usage monitors (SUMs) were on all Envirofit stoves and usually on two three stone fires per household.
*****Each measurement week (weeks 0, 4, 8) involved three 24-hour periods with wood weighing and particulate matter (PM2.5) monitors.
******After we delivered the second Envirofit stove in week 8 we had a shortage of SUMs, so some homes only had a SUM on one three stone fire.

Table 2

Baseline summary statistics and balance of covariates

	Control Mean	Control SD	Treatment Mean	Treatment SD	Difference	p-value	N
Household demographics							
Female respondent (share)	0.68	0.47	0.73	0.45	0.05	0.38	164
Age of respondent	44.06	13.46	40.38	12.29	-3.68	0.14	163
Married (share)	0.78	0.42	0.77	0.43	-0.01	0.85	164
Wife is primary cook (share)	0.94	0.24	0.92	0.28	-0.02	0.60	164
Spouses make decisions jointly (share)	0.57	0.50	0.52	0.50	-0.05	0.52	164
Socioeconomic status							
Earns income (share)	0.92	0.28	0.88	0.33	-0.04	0.56	163
Self employed (share)	0.73	0.45	0.73	0.45	0.00	1.00	164
Year round employment (share)	0.52	0.50	0.49	0.50	-0.04	0.62	164
Identify as subsistence farmers (share)	0.85	0.36	0.85	0.36	0.00	1.00	164
Value of assets (USD)	905.10	1240.82	1158.37	1650.68	253.27	0.08	164
Stove use and fuels							
Number at largest daily meal	6.16	1.95	6.51	2.25	0.35	0.23	163
Always boils drinking water (share)	0.74	0.44	0.72	0.45	-0.02	0.69	164
Firewood primary fuel source (share)	0.94	0.24	0.95	0.22	0.01	0.81	164
Purchased firewood last month (share)	0.34	0.48	0.43	0.50	0.08	0.24	162
Gathered firewood last month (share)	0.82	0.39	0.81	0.39	-0.01	0.97	163
Baseline cooking measurements							
Daily hours cooked on primary three stone fire	7.30	6.75	8.14	7.21	0.84	0.47	118
Daily hours cooked on secondary three stone fire	5.91	6.41	4.51	5.68	-1.41	0.28	99
Daily hours cooked on all three stone fires	12.43	9.71	10.34	8.99	-2.10	0.34	91
Net wood used daily (weight in kg)	9.30	4.10	10.02	4.70	0.73	0.38	153
Average PM2.5 reading, $\mu g/m^3$	414.30	240.84	372.66	228.91	-41.64	0.33	150
Number of households receiving offer	82		82				164

Note: Household data collected at parish wide sales meetings. We adjust standard errors for clustering at the parish level. To minimize the effect of outliers the value of assets and PM2.5 readings are top and bottom coded at 98% and 2% of the distribution while wood use is top coded at 98%. Daily hours cooked on all three stone fires is only calculated if non missing values exist for both the primary and secondary three stone fire. The prices used to calculate asset values are taken from the 2011-12 round of the Uganda Living Standards Measurement Survey (LSMS) published by the World Bank. Values presented are rounded to two decimal places, the value in the difference column is calculated prior to rounding.

Table 3

Envirofit stove use

Variable	Mean	Std. Dev.	Min.	Max.	N
Early buyers have one Envirofit					
Daily hours cooked on primary Envirofit	4.35	3.89	0.02	16.75	188
All buyers have two Envirofits					
Daily hours cooked on primary Envirofit	4.25	3.68	0	16.23	198
Daily hours cooked on secondary Envirofit	2.91	3.5	0	16.93	198
Daily hours cooked on all Envirofits	7.17	4.79	0.26	24.59	198

Note: This table only includes data from weeks with a kitchen performance test when households had one or two Envirofits.

Table 4

Effect of the Envirofit on daily wood used for cooking

Dependent variable = kg. of wood used daily

VARIABLES	(1) OLS	(2) FE
Treatment	0.72	
	(0.72)	
Early buyers have one Envirofit	-1.86***	-1.73***
	(0.60)	(0.56)
All buyers have two Envirofits	-2.48***	-2.48***
	(0.68)	(0.66)
Treatment x Early buyers have one Envirofit	-0.95	-1.08*
	(0.85)	(0.56)
Treatment x All buyers have two Envirofits	-0.46	-0.55
	(0.88)	(0.59)
Constant	12.40***	
	(0.46)	
Observations	1,116	1,116
R-squared	0.15	0.42
Number of household fixed effects		163

Standard errors clustered at parish-wave level in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Wood weights are top coded at 98%. OLS regressions include parish fixed effects.

Table 5

Effect of the Envirofit on daily PM concentrations

Dependent variable = natural log daily PM concentrations

VARIABLES	(1) OLS	(2) FE
Treatment	-0.02	
	(0.03)	
Early buyers have one Envirofit	0.12**	0.12**
	(0.05)	(0.05)
All buyers have two Envirofits	-0.10**	-0.10*
	(0.04)	(0.05)
Treatment x Early buyers have one Envirofit	-0.13*	-0.12**
	(0.07)	(0.06)
Treatment x All buyers have two Envirofits	-0.02	-0.02
	(0.06)	(0.06)
Constant	6.57***	
	(0.07)	
Observations	1,242	1,242
R-squared	0.87	0.92
Number of household fixed effects		164

Standard errors clustered at parish-wave level in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: OLS regression includes parish fixed effects and all regressions include PM monitor fixed effects. PM2.5 readings are top and bottom coded at 98% and 2% of the distribution prior to taking the natural log.

Table 6
Effect of the Envirofit on daily hours cooked on three stone fires - all weeks

	Dependent variable = daily hours cooked on all (cols. 1 and 2), primary (cols. 3 and 4), or secondary (cols. 5 and 6) three stone fire(s)					
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	OLS	FE	OLS	FE	OLS	FE
Treatment	-2.58		0.26		-1.86	
	(2.47)		(1.36)		(1.22)	
Weeks 1-4 (Early buyers have one Envirofit)	1.80	1.96**	1.28	1.49*	0.82	1.22***
	(1.79)	(0.83)	(1.00)	(0.84)	(0.82)	(0.32)
Weeks 5-8 (All buyers have one Envirofit)	-2.72	-3.09***	0.34	0.42	-0.73	-1.04**
	(1.82)	(0.95)	(1.19)	(0.88)	(0.90)	(0.42)
Weeks 9-14 (All buyers have two Envirofits)	-3.61*	-5.15***	-0.45	-0.38	-0.13	-0.85
	(2.08)	(1.53)	(1.15)	(0.91)	(0.94)	(0.62)
Treatment x Early buyers have one Envirofit	-3.16	-3.73***	-3.33**	-3.68***	0.15	-0.58
	(2.67)	(0.74)	(1.60)	(1.12)	(1.37)	(0.48)
Treatment x All buyers have one Envirofit	1.83	2.89***	-1.91	-1.77	2.96**	3.07***
	(2.78)	(1.05)	(1.86)	(1.09)	(1.35)	(0.78)
Treatment x All buyers have two Envirofits	-0.29	0.73	-1.47	-1.03	2.66	1.19
	(3.18)	(1.75)	(1.96)	(1.25)	(1.68)	(1.07)
Constant	14.39***		5.63***		6.27***	
	(1.76)		(0.92)		(0.92)	
Observations	8,595	8,595	13,890	13,890	8,056	8,056
R-squared	0.13	0.58	0.10	0.45	0.08	0.52
Number of household fixed effects		144		160		146

Standard errors clustered at parish-wave level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Data includes all weeks that temperature sensors were on stoves. OLS regressions include parish fixed effects.

Table 7

Effect of the Envirofit on daily hours cooked on three stone fires - measurement weeks

Dependent variable = daily hours cooked on all (cols. 1 and 2), primary (cols. 3 and 4), or secondary (cols. 5 and 6) three stone fire(s)

VARIABLES	(1) OLS	(2) FE	(3) OLS	(4) FE	(5) OLS	(6) FE
Treatment	-1.93		0.78		-0.91	
	(2.00)		(1.01)		(1.14)	
Early buyers have one Envirofit	4.35**	2.75	2.56**	3.77***	2.17**	1.55
	(1.93)	(1.95)	(1.16)	(1.01)	(0.86)	(0.94)
All buyers have two Envirofits	-3.56	-10.20**	-1.49	-0.86	1.06	0.94
	(2.85)	(3.81)	(1.19)	(1.34)	(1.62)	(2.40)
Treatment x Early buyers have one Envirofit	-7.41***	-6.36***	-6.56***	-7.79***	-1.09	-1.07
	(2.52)	(1.63)	(1.57)	(1.17)	(1.49)	(0.99)
Treatment x All buyers have two Envirofits	-3.16	3.38	-2.42	-2.53	1.71	0.30
	(3.71)	(4.71)	(1.83)	(1.74)	(3.38)	(3.75)
Constant	12.36***		5.06***		6.73***	
	(1.62)		(0.94)		(0.79)	
Observations	571	571	941	941	555	555
R-squared	0.24	0.73	0.18	0.60	0.13	0.73
Number of household fixed effects		129		155		133

Standard errors clustered at parish-wave level in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: This table only includes data from weeks with a kitchen performance test. OLS regressions include parish fixed effects.

Table 8

Effect of the Envirofit on daily hours cooked on Envirofit(s) - all weeks

Dependent variable = daily hours cooked on all (cols. 1 and 2), primary (cols. 3 and 4), or secondary (col. 5) Envirofit(s)

VARIABLES	(1) OLS	(2) FE	(3) OLS	(4) FE	(5) OLS
Treatment	0.44		0.44		0.07
	(0.35)		(0.55)		(0.35)
Weeks 5-8 (All buyers have one Envirofit)	-0.17	0.05	-0.09	-0.02	
	(0.27)	(0.26)	(0.25)	(0.21)	
Weeks 9-14 (All buyers have two Envirofits)	1.90***	2.24***	0.08	0.04	
	(0.56)	(0.54)	(0.50)	(0.33)	
Treatment x All buyers have two Envirofits	-0.76	-0.90	-0.22	-0.22	
	(0.72)	(0.56)	(0.51)	(0.29)	
Constant	1.59***		1.53**		2.16***
	(0.43)		(0.64)		(0.10)
Observations	6,853	6,853	8,923	8,923	2,957
R-squared	0.12	0.47	0.09	0.41	0.10
Number of household fixed effects		130		152	

Standard errors clustered at parish-wave level in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Data includes all weeks that temperature sensors were on stoves. OLS regressions include parish fixed effects. The constant in column (1) corresponds to the period when early buyers owned one Envirofit.

Table 9

Effect of the Envirofit on daily hours cooked on Envirofit(s) - measurement weeks

Dependent variable = daily hours cooked on all (cols. 1 and 2), primary (cols. 3 and 4), or secondary (col. 5) Envirofit(s)

	(1)	(2)	(3)	(4)	(5)
VARIABLES	OLS	FE	OLS	FE	OLS
Treatment	-0.01		0.21		0.65
	(0.77)		(0.66)		(0.48)
All buyers have two Envirofits	2.71***	3.08***	0.10	-0.36	
	(0.65)	(0.81)	(0.54)	(0.57)	
Constant	3.97***		3.75***		3.00***
	(0.77)		(0.66)		(0.14)
Observations	390	390	482	482	256
R-squared	0.16	0.66	0.05	0.57	0.12
Number of household fixed effects		105		129	

Standard errors clustered at parish-wave level in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: This table only includes data from weeks with a kitchen performance tests. At midline treatment households owned one Envirofit and at endline all households owned two Envirofits. OLS regressions include parish fixed effects. The constant in column (1) corresponds to the period when early buyers owned one Envirofit.

Table 10

Adjustments for Hawthorne effect

	Change in TSF Hours (hr/day)	TSF wood usage (kg/hr)	Change in ENV Hours (hr/day)	ENV wood usage (kg/hr)	Adjustment for Wood (kg/day)
Midline (Early Buyers)	2.63	0.64	-2.38	0.32	0.92
Endline (All Buyers)	5.05	0.64	-3.06	0.32	2.25

	Change in TSF Hours (hr/day)	TSF PM2.5 (μg/m^3 per hr)	Change in ENV Hours (hr/day)	ENV PM2.5 (μg/m^3 per hr)	Adjustment for PM2.5 (μg/m^3 per day
Midline (Early Buyers)	2.63	32.95	-2.38	16.08	48.39
Endline (All Buyers)	5.05	32.95	-3.06	16.08	117.19

Note: Stove users used three stone fires less and Envirofit stoves more when observers were present, when observers departed they reveresed these changes (Simons et al. (2017)). Therefore, to adjust for this observer induced bechavior, we calculate the change in TSF hours per day as the difference in the coefficients when estimating the effect of the introduction of Envirofit(s) on TSF use only in the measurement week compared to all weeks (difference of coefficients between Table 6 and 7). The change in ENV hours per day is calculated as the difference in the coefficients when estimating the effect of the introduction of Enviroift(s) on ENV use only in the measurement week compared to all week (difference of coefficients between Table 8 and 9). Three stone fire wood (PM2.5) usage per hour calculated during first kitchen performance test when no one owned an Envirofit. Envirofit wood (PM2.5) usage per hour calculated using the laboratory results shown in the Emission and Performance Report (Figure 2) because we do not have any periods in our experimental setting when households only had Envirofits.

Table 11

Estimates of Wood Use and PM concentrations after Hawthorne Effect Adjustment

	Baseline Amount (kg/day)	Unadjusted Change (kg/day)	Unadjusted Change (%)	Adjustment (kg/day)	Adjusted Change (kg/day)	Adjusted Change (%)
Midline (Early Buyers)	9.30	-1.08	-11.6%	0.92	-0.16	-1.7%
Endline (All Buyers)	9.30	-2.48	-26.7%	2.25	0.48	-2.5%

	Baseline Amount ($\mu g/m^3$ per day)	Unadjusted Change ($\mu g/m^3$ per day)	Unadjusted Change (%)	Adjustment ($\mu g/m^3$ per day)	Adjusted Change ($\mu g/m^3$ per day)	Adjusted Change (%)
Midline (Early Buyers)	414.30	-49.72	-12.0%	48.39	-1.33	-0.33%
Endline (All Buyers)	414.30	-41.43	-10.0%	117.19	75.76	18.3%

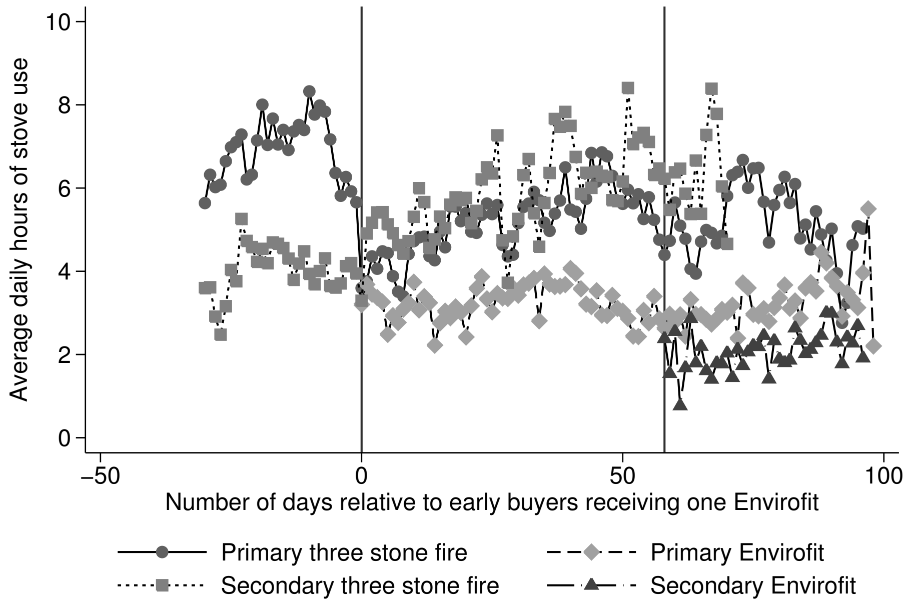
Note: Unadjusted estimates of the change in wood usage come from Table 4. Unadjusted estimates of the change in PM2.5 come from Table 5. The adjustments are calculated in Table 10. Calculations for the adjusted changes are based on Equations 3 and 4. Baseline amounts come from Table 2.

Table 12

Long term usage study: unannounced home visit 3.5 years after initial Envirofits delivered

	N	%
Someone home for unannounced long term usage study	137	100.0%
Actively cooking in moment when enumerators arrived	66	100.0%
-among those, cooking on three stone fire only	52	78.8%
-among those, cooking on Envirofit only	6	9.1%
-among those, cooking on other (mud/charcoal) stove	8	12.1%
Among all households not using Envirofit when enumerators arrived, enumerators asked to see primary Envirofit stove for signs of use	131	100.0%
-primary Envirofit with obvious signs of use	85	64.9%
-primary Envirofit stored and clearly not being used	22	16.8%
-primary Envirofit stored and in perfect condition (basically never used)	3	2.3%
-primary Envirofit damaged and disposed of	11	8.4%
-primary Envirofit given away (condition unknown)	10	7.6%
Among all households that stated they received two Envirofits, enumerators asked to see secondary Envirofit stove for signs of use	129	100.0%
-secondary Envirofit with obvious signs of use	32	24.8%
-secondary Envirofit stored and clearly not being used	14	10.9%
-secondary Envirofit stored and in perfect condition (basically never used)	12	9.3%
-secondary Envirofit damaged and disposed of	21	16.3%
-secondary Envirofit given away (condition unknown)	49	38.0%
Asked: "Do you still use the Envirofit stove?"	137	100.0%
-"I still use both Envirofits"	31	22.6%
-"I still use only one Envirofit"	69	50.4%
-"I have stopped using Envirofits"	37	27.0%
Asked: "If you bought a new stove today, would you purchase an Envirofit?"	137	100.0%
-Yes	108	78.8%
-No	21	15.3%
-Unsure or no response	8	5.8%

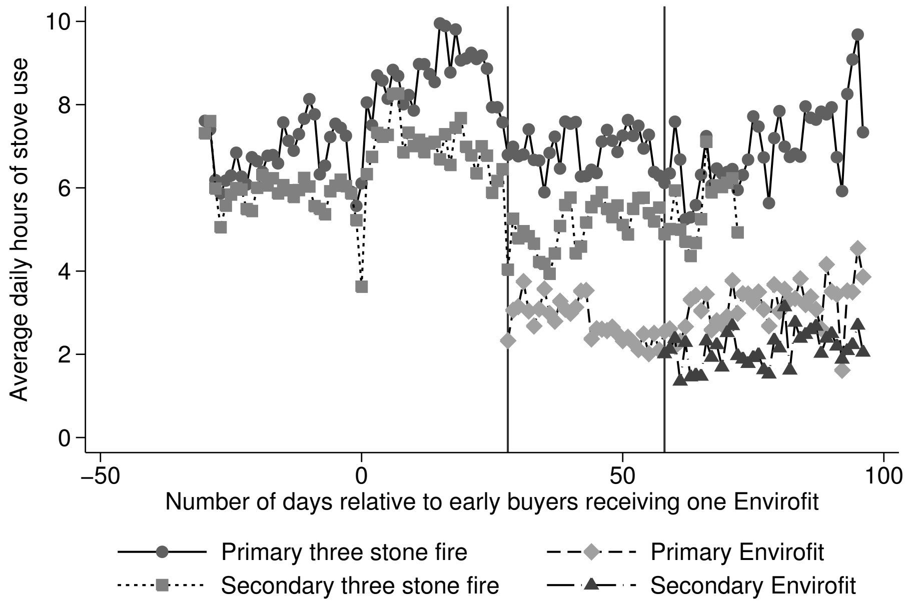
Figure A1
Daily stove use of early buyers



Note: Vertical lines designate when households received their first and second Envirofits.

Figure A2

Daily stove use of late buyers



Note: Vertical lines designate when households received their first and second Envirofits.

Figure A3

SUM holder designed to encase the stove use monitor to protect it from malfunctions when exceeding temperatures of 85 degrees Celsius

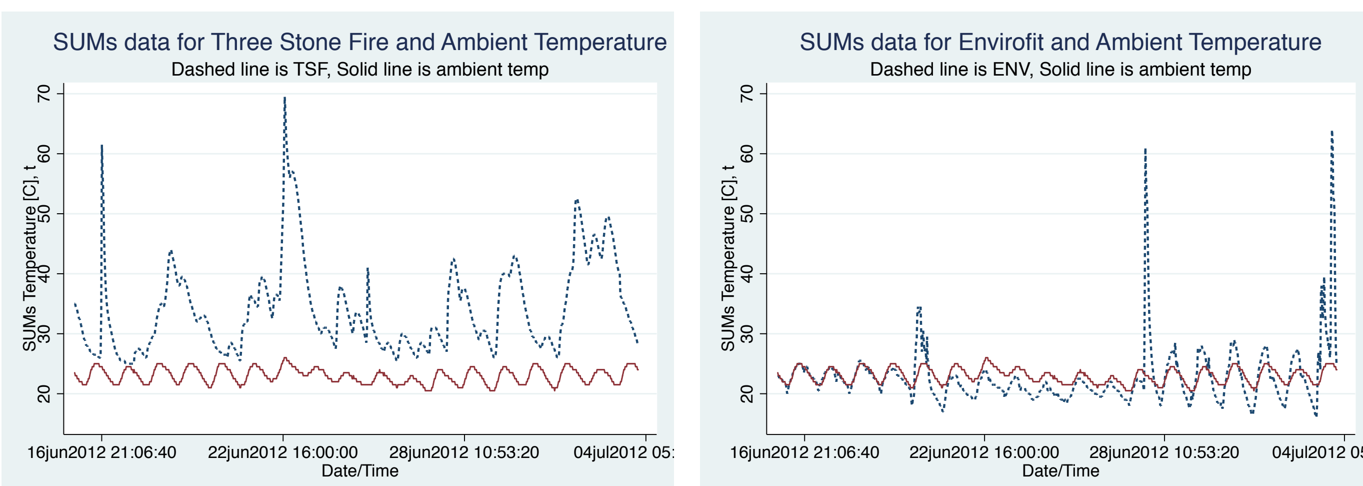
Figure A4

Arrows mark the placement of SUMs on three stone fire and Envirofit

(a) Three Stone Fire (b) Envirofit

Figure A5

Example of household level SUMs temperature data in same household at same times



(a) Three Stone Fire (b) Envirofit

Table A1

SUMs Device Attrition: linear probability model of missing SUMs data

	TSF1 (1)	ENV1 (2)	ENV2 (3)	TSF1 (4)	ENV1 (5)	ENV2 (6)	TSF1 (7)	ENV1 (8)	ENV2 (9)	TSF1 (10)	ENV1 (11)	ENV2 (12)	TSF1 (13)	ENV1 (14)	ENV2 (15)
Daily household wood use (kg)	-0.02	-0.01	-0.01												
	(0.01)	(0.02)	(0.01)												
People cooked for daily				-0.04*	-0.00	-0.00									
				(0.02)	(0.02)	(0.02)									
Meals cooked daily							-0.09	-0.07	0.04						
							(0.06)	(0.05)	(0.06)						
Meals cooked *matooke* daily										0.02	0.10	0.09			
										(0.09)	(0.08)	(0.13)			
Age of cook													-0.00	0.00	-0.00
													(0.00)	(0.00)	(0.00)
Observations	147	142	137	153	146	141	153	146	141	153	146	141	153	147	141
R-squared	0.02	0.00	0.01	0.03	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00
Parish clusters	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14

Standard errors clustered at parish level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The dependant variable is a 0/1 for whether we have the SUMs temperature data during the endline for a specified stove type (TSF1 = primary three stone fire, ENV1 = first Envirofit, ENV2 = second Envirofit). The overall sample only includes stoves that we placed a SUMs device on during the endline. The daily wood weights and counts of cooking practices are averaged across the KPT measurement week. To account for possible correlation in how data was collected by the measurement team, we cluster standard errors at the parish level because the KPT measurement team spent a week at a time in a given parish.